EVOLUTION OF ANGIOSPERM POLLEN. 1. INTRODUCTION[1]

Alexandra H. Wortley,[2] Hong Wang,[3] Lu Lu,[3] De-zhu Li,[3] and Stephen Blackmore[2]

Abstract

This paper is the first in a series that documents the diversity, distribution, and evolution of palynological characters across angiosperms in a contemporary phylogenetic context, using modern optimization methods. The objectives of the series are: (1) to describe the diversity of pollen morphologies across the angiosperms; (2) to estimate ancestral palynological character states, diagnostic characters, and synapomorphies for monophyletic groups; (3) to highlight and interpret inferred patterns and processes of evolution in palynological characters; and (4) to provide a framework for the placement of enigmatic taxa (including fossil taxa) based on pollen morphology. This first paper examines the methods available for such a study and presents an overview of palynological characters across angiosperms as a whole. Using a well-supported, recent, molecular phylogeny, we consider the effects of coding strategy, method of optimization, and starting tree topology upon inference of trait evolution. Coding strategy and optimization method had significant effects upon inferred ancestral character states, the latter probably due to the different evolutionary models applied. Phylogenetic topology had little effect upon inferred ancestral character states, because the uncertainty in topology at this level involved only nodes where few character state changes occurred. Several palynological characters showed consistent, structured patterns in the context of phylogeny: angiosperms are distinguished from other seed plants by character states including supratectal elements echinate and less than 1 μm in size, and infratectum structure columellate; eudicots, as recognized in previous studies, may be defined by globose, isopolar, radially symmetrical grains with three equatorial apertures. We present a framework for the remainder of the series, in which the angiosperms are divided into nine monophyletic and paraphyletic groups each having a similar level of pollen variability, and a set of recommendations for the analysis of these groups. The series will provide a reference for future palynological and systematic studies and an approach that may be replicated for other character sets.

Key words: Ancestral character states, diagnostic characters, evolution, flowering plants, optimization, palynology, phylogeny, synapomorphy, trait evolution.

There is a long and illustrious history of studying palynological characters in a broad evolutionary context (e.g., Wodehouse, 1935; Erdtman, 1943, 1944, 1945, 1946, 1948, 1952; Walker & Doyle, 1975; Nowicke & Skvarla, 1979). The remarkable potential of palynological characters in systematic classification, recently discussed by Blackmore (2000), lies in the fact that pollen grains contain a remarkable degree of information for their minute size, which is in turn due to a number of factors including their highly resistant (almost indestructible) sporopollenin wall, great abundance, and incredible, heritable, variety of form (Blackmore, 2007). The importance of pollen characters in systematics has been highlighted through the recognition of a monophyletic tricolpate clade within angiosperms (Donoghue & Doyle, 1989), now widely known as the eudicots (e.g., Doyle & Hotton, 1991; APG, 1998; APG II, 2003; APG III, 2009). It is likely that further synapomorphic palynological characters remain to be discovered through our growing understanding of both plant phylogeny and pollen development and structure.

The fact that pollen characters are both variable and characteristic of taxa was perhaps first noted by Grew (1682). However, it was more than a century later that Brown (1811) recognized they might therefore be useful for classification (Fig. 1A). The first attempt at a full classification of the flowering plants based on pollen grains was made by von Mohl (1835; Fig. 1B). Von Mohl's work seems to have been largely overlooked until it was rediscovered by

[1] The authors gratefully acknowledge Robert K. Jansen and colleagues for permission to use their plastid genome phylogeny, and Mark Chase for useful discussions on phylogenies and the Angiosperm Phylogeny Group III classification. We also thank the herbaria of BM, E, K, KUN, and US for their help with providing pollen samples, the late John Skvarla for provision of some SEM micrographs, and Hervé Sauquet for supplying the original for Figure 1F. AHW acknowledges the Sibbald Trust at the Royal Botanic Garden Edinburgh for funding to visit China during the project. This work was supported by grants from the National Natural Science Foundation of China (Grant number 31270272), Major International Joint Research Project of the National Natural Science Foundation of China (Grant number 31320103919), and the open funds of the Key Laboratory of Biodiversity and Biogeography, Kunming Institute of Botany, Chinese Academy of Sciences (Grant number KLBB201202).

[2] Royal Botanic Garden Edinburgh, 20a Inverleith Row, Edinburgh EH3 5LR, United Kingdom. Author for correspondence: a.wortley@rbge.ac.uk

[3] Laboratory of Biodiversity and Biogeography, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan, People's Republic of China. wanghong@kib.ac.cn.

doi: 10.3417/2012047

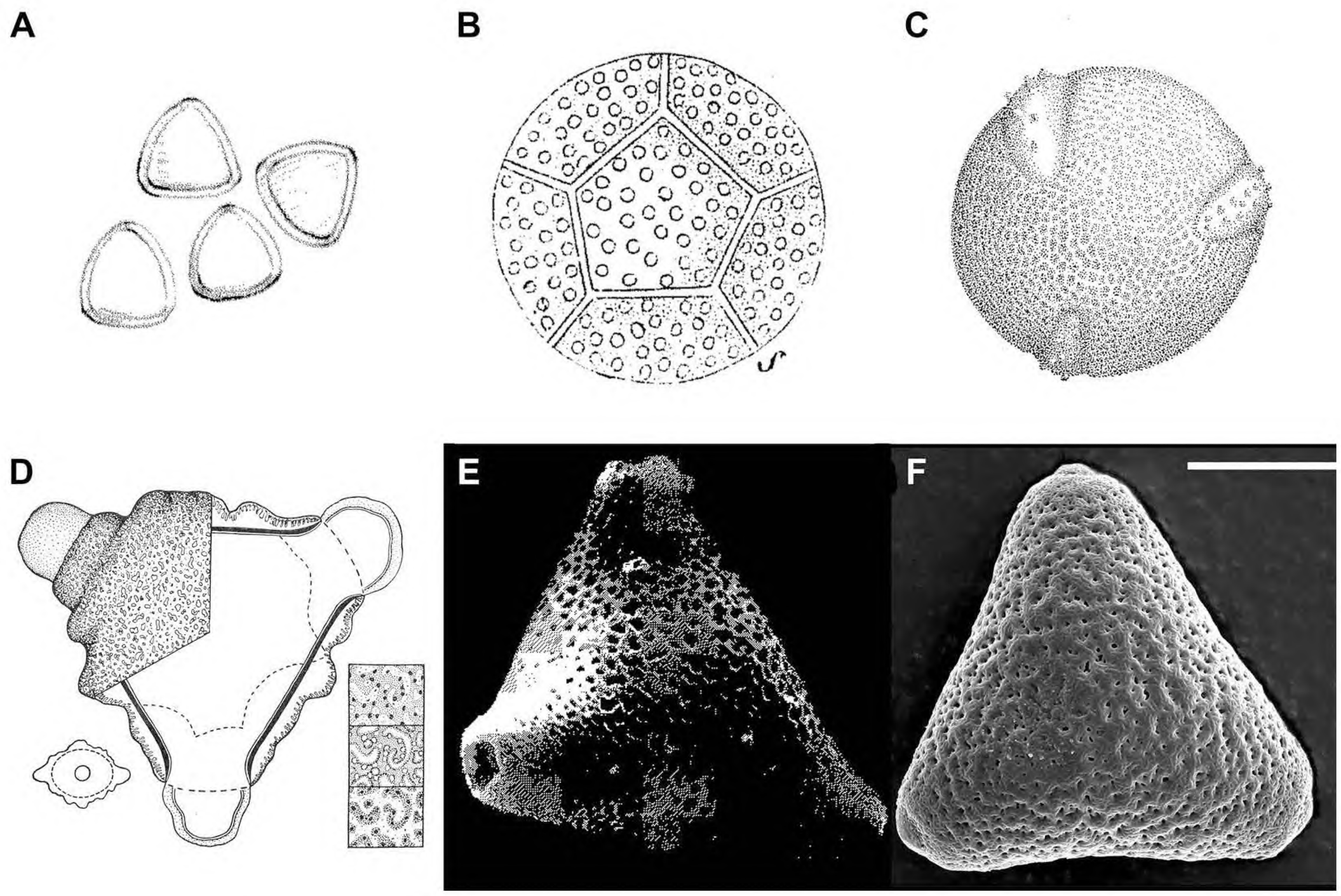


Figure 1. Selected images from the history of palynological study, using examples from the Proteaceae and other basal eudicot groups. —A. *Knightia excelsa* R. Br. (Proteaceae), "pollen plurimum auctum" (mature pollen), reproduced from fig. 7 of Brown's 1811 thesis, *On the Proteaceae of Jussieu*. —B. *Fumaria spicata* L. (Papaveraceae), "graine mouille" (hydrated grain), reproduced from fig. 24 of von Mohl's 1835 work, "Sur la structure et les formes des graines de pollen." —C. *Platanus occidentalis* L. (Platanaceae), polar view, reproduced from plate IX, fig. 5, of Wodehouse's 1935 work, *Pollen Grains: Their Structure, Identification and Significance in Science and Medicine*. —D. *Grevillea bipinnatifida* R. Br. (Proteaceae), fig. 208, reproduced with permission of the Palynological Laboratory, Swedish Museum of Natural History, Stockholm, from the 1972 corrected reprint edition of Erdtman's 1952 epic, *Pollen Morphology and Plant Taxonomy*. —E. *Euplassa inaequalis* Engl. (Proteaceae), triporate pollen, polar view, an early SEM micrograph (fig. 3B) reproduced with kind permission of Missouri Botanical Garden Press from Walker and Doyle's 1975 study, "The bases of angiosperm phylogeny: Palynology." —F. *Sorocephalus crassifolius* Hutch. (Proteaceae), whole grain, polar view, reproduced with kind permission of the American Society for Plant Taxonomy from fig. 15A of Sauquet and Cantrill's 2007 comprehensive evolutionary study, "Pollen diversity and evolution in Proteoideae (Proteales: Proteaceae)."

Wodehouse (1935). Wodehouse himself published a series of beautifully illustrated works on the structure of pollen grains and their use in classification (e.g., Fig. 1C), focusing on Asteraceae (Wodehouse, 1926, 1928a, 1928c, 1929a, 1929b, 1930, 1936). Shortly after this, Erdtman (1943) produced the first of his many surveys of pollen and spores throughout the land plants. His work, *Pollen Morphology and Plant Taxonomy* (Erdtman, 1952), remains the most comprehensive reference guide to angiosperm pollen. It not only describes and illustrates pollen grain structure in great detail (e.g., Fig. 1D), but also draws parallels between families based on palynological characteristics. Sadly, only a small proportion of Erdtman's work has been published, with descriptions and illustrations of a great range of taxa remaining only in the archives of the Swedish Museum of Natural History in Stockholm.

The 1970s saw a gradual progression from simple descriptions of pollen morphology to hypotheses of evolutionary change in pollen characters. Muller (1970) published an evolutionary scenario showing relationships between pollen grain types linked to plant evolution as then understood. Walker and Doyle (1975), using some of the first SEM micrographs (e.g., Fig. 1E), continued in this evolutionary vein at the Bases of Angiosperm Phylogeny symposium (Walker, 1975) at the 24th annual meeting of the American Society of Plant Taxonomists, presenting similar diagrams to Muller (1970) and highlighting evolutionary trends correlated with relative advancement of taxa for a wide range of pollen morphological characters (Walker & Doyle, 1975). Later, at the Evolutionary Significance of the Exine meeting (Ferguson & Muller, 1976), Van Campo (1976) explored trends of palynological evolution in

various angiosperm taxa. Shortly afterward, Nowicke and Skvarla (1979) published a study of pollen morphology and development across 650 species focusing on the Centrospermae (roughly corresponding to today's Caryophyllales). Their study determined diagnostic characters for higher-order groupings, albeit without explicit reference to phylogenetic relationships (Nowicke & Skvarla, 1979). These evolutionary approaches were complemented by Muller's ongoing documentation of angiosperm pollen in the fossil record (Muller, 1970, 1981, 1984). In sum, advances in phylogenetic, palynological, and paleobotanical research, including the relatively new techniques of SEM and transmission electron microscopy (TEM), provided a solid foundation for future interpretations of pollen and spore morphology in a phylogenetic and ontogenetic context, provided by concurrent progress in the visualization of development and numerical (cladistic) methods for phylogenetic analysis (for discussion see Funk & Stuessy, 1978).

Together, these precipitated in the early 1980s a burst of cladistic analyses of pollen characters (alone or in combination with other morphological characters) at familial and lower levels. Early examples include the works of Blackmore (1982) on the Scorzonerinae (Asteraceae), Lee and Park (1982; Oleaceae), Donoghue (1983; *Viburnum* L. [Caprifoliaceae]), Kress and Stone (1983; Heliconiaceae), Blackmore and Cannon (1983; Morinaceae), and Linder (1984; Restionaceae; see Appendix 1 for more recent examples). In such studies, morphological character states distributed across groups in a phylogeny are interpreted as homoplasies, synapomorphies, autapomorphies, or plesiomorphies, all of which may also be diagnostic for the purposes of classification. These concepts enable the generation of evolutionary hypotheses regarding ancestral and derived states and the direction and pattern of evolution in the observed traits. It may also be possible to identify correlations between the evolution of palynological character states, life history strategies and syndromes, and evolutionary-ecological factors such as the origin or diversification of particular pollinating animal groups.

Notwithstanding the conceptual and factual advance represented by the first cladistic studies, they were limited in two respects. Firstly, these early studies typically tended to combine the use of palynological characters as evidence for the reconstruction of phylogenetic relationships with their interpretation in the context of the resultant phylogeny (for an early exception see Jansen et al., 1991). In the absence of an independent phylogenetic estimate, for instance based on molecular genetic data, the interpretation of characters as ancestral or derived, synapomorphy or homoplasy can thus be seen as a circular argument (for discussion see de Queiroz, 1996); others argue that recognition of characters as synapomorphic or homoplastic is intrinsically, and rightly, bound up in the process of phylogenetic estimation (most famously Patterson, 1982). Secondly, perhaps due to the large amount of data required, analysis has been limited to the family level and below. Even had it been possible to explore further in the taxonomic hierarchy, the results would have been erroneous because understanding of angiosperm phylogeny at the time (e.g., Takhtajan, 1980; Cronquist, 1981; Dahlgren, 1988; Thorne, 1992) differed considerably from present-day interpretations. For instance, Walker and Doyle (1975) utilized the broadly phenetic classifications of Cronquist (1988) and Takhtajan (1980), which significantly influenced their inference of trait evolution.

Since the 1990s, there has been a concerted shift toward classifications that prioritize monophyly. For flowering plants, the widely accepted Angiosperm Phylogeny Group (APG) classifications are based largely on molecular phylogenetic studies conducted over the last two decades (APG, 1998; APG II, 2003; APG III, 2009). The first APG classification (APG, 1998) recognized 462 families in 40 putatively monophyletic orders and a few higher groups. Five years on, subsequent changes in family circumscription and ordinal definition were incorporated (APG II, 2003), giving a classification with 457 families in 45 orders. The most recent edition (APG III, 2009) consolidates these into 415 families among an increased 58 orders. Overall, the APG classifications have proven extremely stable. In the most recent, the authors note that "we do not see the APG classification as continuing to mutate for the indefinite future. . . . We hope the classification . . . will not need much further change" (APG III, 2009: 106).

Enabled by such advances in phylogenetic classification, pollen development, electron microscopy (Blackmore, 2007), and methods for optimization of character states upon phylogenies (e.g., Cunningham et al., 1998; Mooers & Schluter, 1999; Omland, 1999; Pagel, 1999a, 1999b; Huelsenbeck et al., 2000; Huelsenbeck & Bollback, 2001; Nielsen, 2002; Pagel et al., 2004; Ronquist, 2004; Vanderpoorten & Goffinet, 2006; Ekman et al., 2008), we can now re-evaluate the evolution of palynological characters across the flowering plants within a more explicit phylogenetic, evolutionary, and ontogenetic context. Similar angiosperm-wide studies have suc-

cessfully been carried out on the evolution of floral symmetry (Jabbour et al., 2009) and secondary chemicals (Esteban et al., 2009). For palynological characters, evolutionary studies have been conducted in the Proteaceae (Sauquet & Cantrill, 2007, an SEM micrograph from which is shown in Fig. 1F) and Asteraceae (Blackmore et al., 2009) which, with ca. 20,000 species (Funk et al., 2005), comprises 7%–10% of flowering plants. Both Sauquet and Cantrill (2007) and Blackmore et al. (2009) demonstrated that in the context of robust phylogenies, pollen characters can provide diagnostic and synapomorphic characters at a range of hierarchical levels, enabling the formulation of adaptive hypotheses. Furthermore, the recovery of ancestral character state combinations may better enable the placement of fossil pollen grains in a phylogenetic context, facilitating greater understanding of the date and location of key events in angiosperm evolution.

In a series of papers, of which this is the first, we will survey the evolution of pollen morphologies throughout angiosperms, based on the most recent and reliable phylogenetic trees and incorporating an extensive literature review and new palynological data. Our objectives are: (1) to describe and document the diversity of pollen morphologies across the angiosperms; (2) to estimate ancestral palynological character states for monophyletic groups and thereby to identify diagnostic characters and synapomorphies; (3) to highlight and interpret inferred patterns and processes of evolution in palynological characters; and (4) to provide a framework for the phylogenetic placement of enigmatic taxa based on pollen morphology (this is especially relevant to extinct taxa since pollen is often particularly well preserved in the fossil record). This first paper compares the available methods for investigating trait evolution in a phylogenetic context and provides an overview analysis of angiosperm palynological characters and their evolution at a broad scale. We hope this will stimulate similar review and analysis of other classes of morphological and anatomical characters.

The Influence of Methodological Factors in the Analysis of Trait Evolution

Three fundamental methodological factors may influence the inference of ancestral morphological character states upon phylogenetic trees: coding strategy, analytical (optimization) method, and starting tree topology (i.e., phylogenetic uncertainty).

Coding morphological characters becomes nontrivial for higher (supraspecific) taxa when more than one character state is found therein. Three main coding strategies are recognized in phylogeny reconstruction; these may equally be applicable to studies of trait evolution. The ancestral (sensu Bininda-Emonds et al., 1998) or groundplan (sensu Yeates, 1995) method uses fossils, ontogenetic data, or prior phylogenetic analysis to estimate a primitive state in each monophyletic higher taxon. The democratic (Bininda-Emonds et al., 1998) or composite (Brusatte, 2010) method uses data from a sample of taxa within each higher taxon, which is then represented by the most common state found (based on the hypothesis that "common equals primitive"). The exemplar method (Bininda-Emonds et al., 1998) employs one or more representatives of each higher taxon, chosen either for their inferred plesiomorphic status or simply their availability. A fourth possibility, to code higher taxa as polymorphic, is not usually considered suitable for reconstruction of phylogenetic trees but is, however, open to character studies based upon existing phylogenies. Like the democratic method, it uses a sample of taxa from across the higher taxon but does not assume that common equals primitive; the higher taxon is coded for *all* states found in its representatives. We term this the comprehensive method.

The choice of analytical methods for character state optimization has received less attention in the literature than that for reconstructing phylogenies (Pagel, 1997; Cunningham et al., 1998), and their implementation remains controversial (Yang, 2006; Wiens et al., 2007). As with phylogeny reconstruction, both parsimony and model-based methods exist. Maximum parsimony (MP), the most widely used method (Cunningham et al., 1998), estimates ancestral states so as to minimize the number of state changes implied across the phylogeny as a whole. This method has proven informative and further has the advantage that it is insensitive to the density of taxon sampling within a clade for which a trait is fixed (D. Barker, 2011, pers. comm.). However, MP by definition assumes that changes of state are rare, i.e., a maximum of one per branch (Cunningham et al., 1998), which may be a weakness particularly when analyzing characters at higher taxonomic levels. In addition, the chance of a state change occurring is considered the same on every branch regardless of length, and MP does not provide probabilistic estimates of error or support for ancestral states, showing only the most parsimonious even if only marginally so. In contrast, probabilistic, model-based approaches (including maximum likelihood [ML] and Bayesian methods) can incorporate measures of uncertainty and/or probability in ancestral states and, where an appropriate model is selected, make use of branch length information (Cunningham et al.,

1998). However, these methods tend to be sensitive to taxon sampling and to the accuracy of branch lengths (Torices, 2010). Of these, ML techniques optimize ancestral states so as to maximize the probability of the observed states in the terminal taxa, given a single explicit model of evolution (for morphological data this is usually simple, such as a one-parameter Markov [Mk-1] model, in which the single parameter is the rate of change). ML analysis assumes a constant evolutionary rate across the tree, i.e., the likelihood of change along a branch is proportional to its length (Cunningham et al., 1998). For phylogenies based on molecular evidence, this may be problematic if the morphological events of interest are uncoupled from molecular change, such as when morphological change is concentrated during speciation events. The technique sometimes known as empirical Bayesian (EB) inference is considered by some authors to be superior to ML methods (e.g., Yang, 2006). Like ML, it uses an Mk-1 model of transition rates derived from the data, but unlike the former it subsequently derives posterior probability distributions for ancestral states. Fully (hierarchical) Bayesian (HB) techniques have the advantage that they can incorporate multiple alternative base phylogenies, by calculating relative probability for each possible character state at each node on a set of trees sampled from a distribution (usually generated during Bayesian inference of phylogeny). However, some studies have suggested that the advantages of optimization over multiple phylogenies are minimal (Hanson-Smith et al., 2010). Possible disadvantages of Bayesian methods include their dependence on the choice of prior probabilities, which may be somewhat arbitrary. Both ML and Bayesian methods are limited in applicability since they can at present be conducted only on fully resolved phylogenies and datasets without polymorphisms.

Although the starting tree we employed for comparisons of coding and optimization method (Jansen et al., 2007) was the most reliable and suitable phylogeny available for this study, it is not the only possible representation of relationships between angiosperm taxa. Recent studies have found a variety of patterns among some of the earliest branching lineages of the angiosperm tree, and even a thorough analysis of whole genome data, aimed specifically at resolving relationships among basal angiosperms, failed to reach resolution on this issue (Moore et al., 2007). Although the majority of plastid-based studies have agreed on the topology found by Jansen et al. (2007), since topology is crucially important to polarizing character states we take the conservative view of Soltis et al. (2008) that this should not be considered settled until comparisons have been made with nuclear sequence data.

Materials and Methods

Study taxa

Study taxa were determined by the starting phylogeny (Fig. 2), which was taken from the work of Jansen et al. (2007), and comprised 64 taxa in 59 genera (61 angiosperm species in 56 genera plus three gymnosperm outgroups). This phylogeny was selected from among those in the recent literature because it provided broad taxonomic coverage across a wide range of genetic loci, was fully resolved and well supported, featured branch lengths calculated with model-based methods, was based on the entire chloroplast genome (minimizing issues of rapid evolution at certain loci in certain taxa), and is highly congruent with other recent (including subsequent) phylogenetic estimates (e.g., Soltis et al., 2011). Pollen data were, if possible, taken from the same 64 species as sampled in the starting phylogeny. However, where palynological data were not available for these taxa, data were obtained from closely related species or genera (see Appendix 2 for a list of substituted taxa).

Pollen characters

Character state information was obtained from the literature, with taxon-specific palynological studies providing the primary reference where available. We also drew upon the broader studies of Sampson (2000) for magnoliids, Zavada (1983) and Furness and Rudall (2003) for monocots, Punt (1984; Umbelliferae) and Blackmore et al. (2009; Asteraceae), as well as Wodehouse (1935), Erdtman (1952), Walker and Doyle (1975), and Nowicke and Skvarla (1979; cf. Appendix 1). When possible, character state information was taken directly from original published images (light microscopy [LM], SEM, and TEM) and measurements rather than written descriptions, in order to achieve a more consistent description and terminology for character states across all taxa. All observations pertain to mature pollen grains. Differences in treatment and hydration state (for instance, of fresh material compared to that taken from herbarium specimens) can have a significant effect upon the nature of certain characters such as size and shape. Therefore, acetolyzed (Erdtman, 1960) grains were used wherever possible to ensure comparability of all characters both within this series and with the majority of palynological literature.



Figure 2. Angiosperm phylogeny redrawn from a maximum likelihood (ML) analysis of plastid genome data for 64 taxa, provided by Jansen et al. (2007), with branch lengths proportional to number of inferred nucleotide substitutions and taxon names as in the original article. This phylogeny served as the basis for most of the analyses presented here. Modified topologies used to test the effects of phylogenetic uncertainty in optimization of ancestral character states are described in the text. Clades represented by at least two taxa in the phylogeny are labeled if mentioned in the text.

As this is a preliminary study, character selection focused upon characters previously identified as variable and potentially informative at intra-ordinal level, including dispersal unit, polarity, symmetry, shape, size, aperture characters, exine structure and sculpture, and external features such as orbicules and viscin threads (Erdtman, 1952; Van Campo & Lugardon, 1973; Walker & Doyle, 1975; Skvarla et al., 1978; Nowicke & Skvarla, 1979; El-Ghazaly & Chaudhary, 1993; Schols et al., 2001; Merckx et al., 2008). For the purposes of data exploration, an inclusive approach was taken in which as many characters as possible were coded, including those that appeared to be somewhat subjective, poorly characterized, or potentially non-independent (note that non-independence of characters is only a concern for phylogenetic reconstruction, not for investigation of character evolution upon an independently derived phylogeny). For very complex features such as exine sculpture it remains difficult to design a perfect strategy for delimitation of characters and character states.

Palynology has several complex descriptive vocabularies, developed for the efficient portrayal of numerous features. The terminology used here follows that of Punt et al. (2007). One widespread shortcut used to avoid lengthy descriptions and complex terminology is the "pollen type."; Pollen types describe the entire nature of a pollen grain (size, shape, apertures, exine stratification, surface sculpture, etc.) in a single phrase such as type A, or Liguliflorae-type. While this approach is useful for the purpose of identifying isolated, dispersed pollen grains, as in fossil studies, to analyze pollen grain evolution, it is more appropriate to break down pollen types and compound terms into their most basic components, each conveying a character. For example, instead of the compound term ana-zona-sulcate we would describe aperture shape (zonate), orientation (latitudinal), structure (simple), and position (between the equator and distal pole). The palynological characters and character states investigated were as follows, organized in the sequence of Erdtman (1952; cf. Appendix 3 and Fig. 3).

Dispersal unit. The dispersal unit (Character 1) is the arrangement in which pollen grains are found at maturity and dispersal (Punt, 1962), also known as the pollen unit (Walker & Doyle, 1975). States range from monads (free grains; e.g., Fig. 3A) to permanent dyads (fused pairs), tetrads (fused groups of four, in various orientations; Fig. 3B), polyads (fused in defined multiples of four, to a maximum of 64; Fig. 3C), or pollinia (fused in larger, indeterminate numbers, to date only reported for Apocynaceae and Orchidaceae). In some taxa, notably the Cyperaceae, pollen grains are initially borne in tetrads but reduced by abortion to single grains, known as cryptotetrads or pseudomonads (Erdtman, 1952), soon after meiosis; we treat these structures as tetrads.

Polarity and symmetry. Polarity (Character 2) and symmetry (Character 3) are determined with reference to development in the tetrad: each pollen grain has a polar axis running outward from the center of the tetrad, and two poles where this axis meets the surface of the grain, a proximal pole at the center of the tetrad and a distal pole at the outer surface (Erdtman, 1952). Based on this, pollen grains may be apolar (Fig. 3A), isopolar (Fig. 3D), subisopolar, or heteropolar (Fig. 3E). In apolar grains, such as many polyaperturate grains, orientation and polarity are impossible to establish at the free microspore stage. In isopolar grains, the two hemispheres are identical, as in most tri-aperturate, spheroidal, eudicot grains. Subisopolar grains are not entirely symmetrical about the equatorial plane, e.g., due to the addition of viscin threads on one face of the grain. In this paper, we treat subisopolar grains as isopolar, since they are usually the same in fundamental structure. In heteropolar grains, the two polar hemispheres differ distinctly in shape or apertures (Erdtman, 1952).

When viewed from either pole, pollen grains typically have either two (bilateral symmetry; Fig. 3F) or more than two (radial symmetry; Fig. 3G) planes of symmetry running perpendicular to the equator. The former is typical of monocolpate grains, the latter of most tri-zono-aperturate grains (Erdtman, 1952).

Shape. The basic three-dimensional shapes of pollen grains (Character 4) are often categorized as either globose or boat-shaped. Other shapes, such as threadlike, are restricted to a few species with reduced exines. Boat-shaped pollen, found in many grains with a single, colpate, polar aperture, has a short polar axis and unequal equatorial axes (Walker & Doyle, 1975; Fig. 3F). Globose pollen, the most common type, is approximately spheroidal to ellipsoidal (e.g., Fig. 3A, D, E, G–P). These two broad states are included in our matrix in order to assess whether the character may be informative. However, since many grains do not precisely fit such categories, it has become more conventional to describe shape in terms of two outlines: in equatorial and in polar view. In equatorial view, grains may be defined by their shape class (Character 5; Erdtman, 1952), which is defined as the ratio of the lengths of their polar and

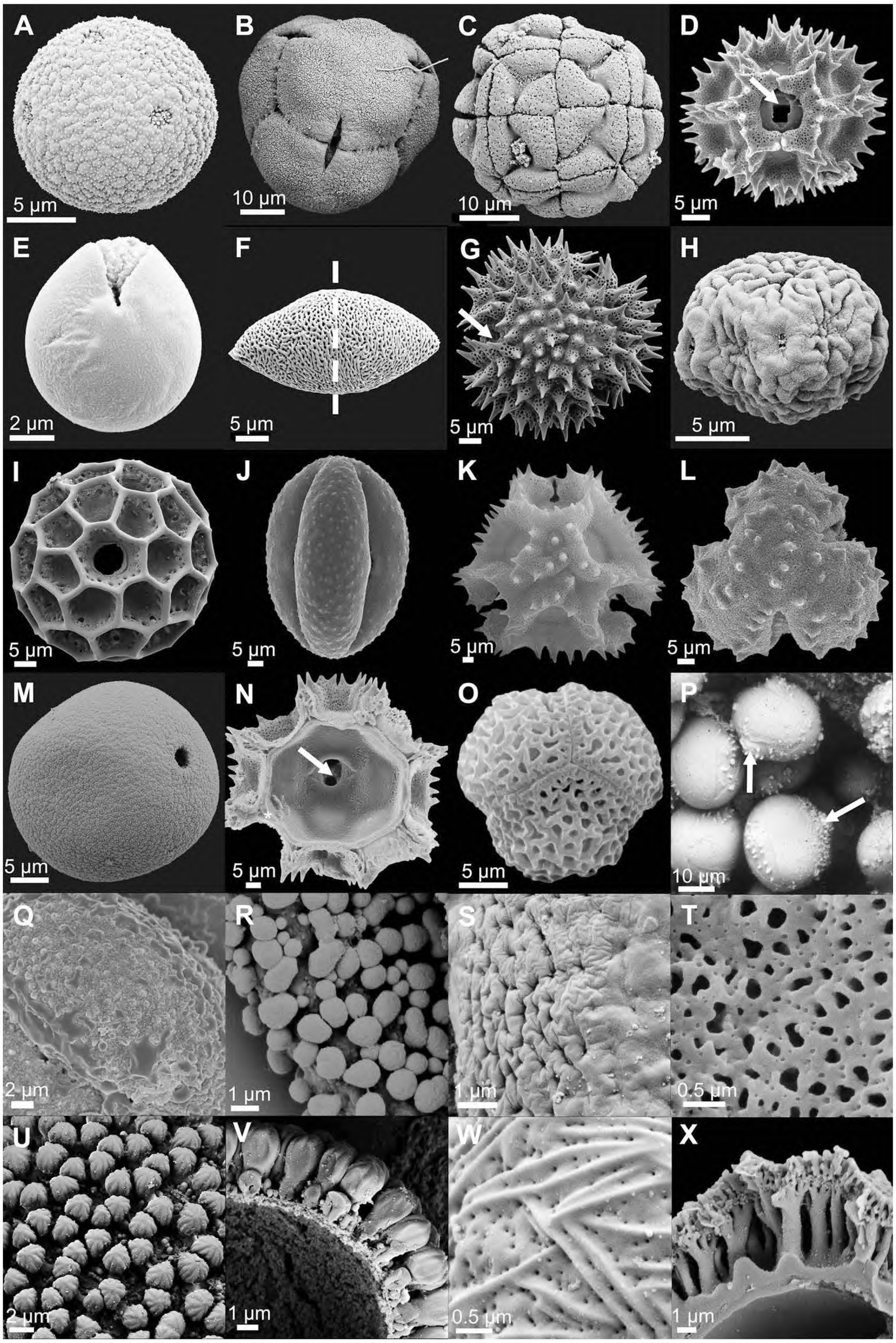


Figure 3. SEMs of pollen grains representative of some of the characters and character states used in this paper. For voucher details see Appendix 6. A–P. Whole grains or clusters of grains. —A. *Plantago psyllium* L. (Plantaginaceae), unknown view, illustrating dispersal unit (Character 1) as monads, pollen grains apolar (Character 2), basic shape (Character 4) globose, apertures (Character 8) many, aperture position (Character 9) global. —B. *Rhododendron wallichii* Hook. f. (Ericaceae), dispersal unit (Character 1) as tetrads. —C. *Acacia nilotica* (L.) Willd. ex Delile (Fabaceae), dispersal unit (Character 1) as polyads. —D. *Crepis napifera* (Franch.) Babc. (Asteraceae), equatorial view, pollen grains isopolar (Character 2), aperture membranes (Character 12;

equatorial axes (P/E). A P/E of less than 0.5 is defined as peroblate, 0.5 to 0.75 as oblate (Fig. 3H), 0.75 to 1.33 as subspheroidal (Fig. 3I), 1.33 to 2.0 as prolate (Fig. 3J), and greater than 2.0 as perprolate (Erdtman, 1952). The widespread subspheroidal class may further be divided into four categories: suboblate (P/E 0.75–0.88), oblate-spheroidal (P/E 0.88–1.0), prolate-spheroidal (P/E 1.0–1.14), and subprolate (P/E 1.14–1.33; Erdtman, 1952). In polar view (amb; note that in heteropolar grains this does not necessarily coincide with the equatorial outline), the outline (Character 6) may be described as circular (Fig. 3I), elliptical, polygonal (Fig. 3K), concave-polygonal (angular with concave sides), or lobate (curved with convex sides separated by indentations; Fig. 3L).

Size. We follow the convention of Erdtman (1952) and Walker and Doyle (1975) in measuring size (Character 7) as the length of the longest axis of each pollen grain. As a continuous character, size can only arbitrarily be divided into states; furthermore, because size may vary with preparation of the grain, it has been suggested that any states should be defined as separated at least by orders of magnitude (Walker & Doyle, 1975). We adopt the following standard classes (Erdtman, 1952; Walker & Doyle, 1975): very small (less than 10 μm), small (10–24 μm), medium (25–49 μm), large (50–99 μm), very large (100–199 μm), and gigantic (greater than 200 μm).

Apertures. Apertures, found in the majority of species, have been recognized as fundamental features of pollen grains (Wodehouse, 1928b; Erdtman, 1952; Walker & Doyle, 1975). They comprise modifications of the pollen wall such as openings, thinnings, or thickenings of the exine or intine for the purposes of interaction with the surrounding substrate, including the emergence of pollen tubes during germination. Erdtman (1952) considered seven aspects of aperture morphology, some of which may be interdependent: number, position, structure, shape, size, nature of the aperture membrane, and presence or absence of an operculum. In this paper, we consider all except aperture size. Note that while Erdtman (1952) described a series of terms indicating both position and shape of apertures, including sulcus (a furrow-shaped aperture situated at the distal pole), porus (a rounded aperture at the equator), and rugus (one of several, globally distributed furrows), for the present analysis we prefer to maintain the two characters (position and shape) as separate, for reasons described above.

The number (Character 8) and position (Character 9) of apertures is correlated with developmental factors during and after meiosis (Blackmore & Crane, 1988; Blackmore et al., 2007). In terms of number, in this study we use states zero, one (Fig. 3E, M), two, three (Fig. 3G, K, L), four to six, seven to 12, and more than 12 (Fig. 3A). In terms of position, apertures are defined as polar (proximal or distal; Fig. 3E), equatorial (Fig. 3G), or global (Fig. 3A).

←

arrowed) smooth, supratectal elements (Character 17) echinate. —E. *Acorus gramineus* Sol. ex Aiton (Acoraceae), equatorial view, pollen grains heteropolar (Character 2), apertures (Character 8) one, aperture position (Character 9) distal. —F. *Dioscorea nipponica* Makino (Dioscoreaceae), equatorial view, symmetry (Character 3) bilateral (dashed line marks single plane of symmetry), basic shape (Character 4) boat-shaped. —G. *Centrapalus pauciflorus* (Willd.) H. Rob. (Asteraceae), polar view, symmetry (Character 3) radial, apertures (Character 8) three, aperture position (Character 9) equatorial, aperture membranes (Character 12; arrowed) granulate, supratectal elements (Character 17) echinate. —H. *Ulmus glabra* Huds. (Ulmaceae), equatorial view, shape class (Character 3) oblate, tectum sculpture (Character 19) rugulate. —I. *Cabobanthus bullulatus* (S. Moore) H. Rob. (Asteraceae), equatorial view, shape class (Character 3) subspheroidal, ectoapertures (Character 11) porate, exine (Character 29) lophate. —J. *Nouelia insignis* Franch. (Asteraceae), equatorial view, shape class (Character 3) prolate, ectoapertures (Character 11) colpate. —K. *Tragopogon longifolius* Heldr. & Sartori (Asteraceae), polar view, outline (Character 6) polygonal, apertures (Character 8) three, supratectal elements (Character 17) echinate. —L. *Oldenburgia paradoxa* Less. (Asteraceae), polar view, outline (Character 6) more or less lobate, apertures (Character 8) three, tectum sculpture (Character 19) perforate. —M. *Poa bulbosa* L. (Poaceae), oblique view, apertures (Character 8) one, aperture structure (Character 10) simple, tectum sculpture (Character 19) areolate. —N. *Scorzonera hispanica* L. (Asteraceae), internal view of aperture, aperture structure (Character 10) compound, endoaperture (character not analyzed due to greater than 50% missing data; arrowed) lalongate, endexine (Character 23; asterisked) present. —O. *Illicium floridanum* J. Ellis (Illiciaceae), polar view, ectoapertures (Character 11) syncolpate (reproduced with kind permission of Elsevier from fig. 8 of Wang et al., 2009b). —P. *Aesculus hippocastanum* L. (Sapindaceae), multiple grains, operculum (Character 13; arrowed) present. Q–X. Details of exine sculpture and structure. —Q. *Dioscorea pyrenaica* Bubani & Bordere ex. Gren. (Dioscoreaceae), supratectal elements (Character 17) gemmate. —R. *Adenocaulon chilense* Less. (Asteraceae), detail of exine surface close to aperture, supratectal elements (Character 17) verrucate. —S. *Lathyrus pratensis* L. (Fabaceae), tectum sculpture (Character 19) fossulate. —T. *Cullumia rigida* DC. (Asteraceae), tectum sculpture (Character 19) reticulate. —U. *Croton argyratus* Blume (Euphorbiaceae), tectum sculpture (Character 19) *Croton*-patterned. —V. *Croton argyratus*, fractured exine in cross-section, tectum sculpture (Character 19; shown in cross-section) *Croton*-patterned. —W. *Dampiera stricta* (Sm.) R. Br. (Goodeniaceae), tectum sculpture (Character 19) striate. —X. *Alfredia cernua* (L.) Cass. (Asteraceae), fractured exine in cross section, infratectum structure (Character 20) columellate.

Furthermore, for polygonal grains, the positioning of the apertures relative to the apices of the grain may be a useful character: in angulaperturate grains the apertures lie at the apices; in planaperturate grains they lie on the faces between the apices.

Aperture structure (Character 10) may be simple (Fig. 3M), comprising a single-layered thinning (ectoaperture) in the outer exine, or compound (Fig. 3N), in which there is also a thinning or opening (endoaperture) in the inner exine. Occasionally, a mesoaperture is present between the ecto- and endoapertures; its presence or absence is treated here as a binary character.

In terms of shape, the ectoaperture (Character 11) may be more or less porate (rounded; Fig. 3I), colpate (furrowed; Fig. 3J), zonate (encircling the grain in a ring), spiraperturate (encircling the grain in a spiral pattern), or syncolpate (with elongated apertures fusing, usually at the poles; Fig. 3O). The endoaperture (not analyzed here due to high levels of missing data) may be circular, endocingulate (encircling the grain in a ring), lalongate (elliptical or rectangular with the longest axis parallel to the equator; Fig. 3N), or lolongate (longest axis perpendicular to the equator).

The thin membrane that usually covers an aperture (Character 12) may be smooth (psilate; Fig. 3D) or beset with granules (Fig. 3G). In certain taxa, such as many Poaceae, the aperture membranes are conspicuously thickened with an island of ectexinous material toward the center, known as an operculum (Character 13; Fig. 3P). The apertures may also each be surrounded by an annulus (Character 14), a conspicuous thickening of the exine, which may take several forms: aspidate (outer exine thickened and protruding outward around the aperture), costate (inner exine, endexine, or foot layer thickened and protruding inward into the cytoplasm of the grain), or vestibulate (annulus subtended by a cavity).

Exine structure and sculpture. Most pollen grains are surrounded by an acetolysis-resistant sporopollenin exine (exceptions include many Zingiberales, some aquatic angiosperms, and a few Lauraceae). The exine is usually differentiated (Character 15) into distinct layers, for which there are two alternative descriptive systems. Erdtman (1952) defined two layers based on morphological observations only: an outer sexine and an inner nexine, both of which may be further subdivided into numbered layers. However, when comparing pollen grains with differing numbers of layers, this frequently results in non-homologous layers being given the same number. For this reason, we adopt the less ambiguous system of Fægri (1956) and Fægri and Iversen (1989). These authors defined terms ectexine (or ektexine) and endexine based on the developmental origins and staining properties of the layers. The endexine is usually single-layered, whereas the ectexine is often multi-layered, comprising a tectum and infratectum (which together correspond to the sexine of Erdtman [1952]) and foot layer (which, together with the endexine, comprises Erdtman's nexine). On the surface of the tectum there may also be supratectal elements.

Working inward from the external surface of the pollen grain, we first characterize the nature of the supratectal elements, if present (Character 16). In terms of shape (Character 17), they may be pilate (approximately cylindrical, greater in height than diameter), echinate (pointed, broader at base than at tip; Fig. 3D, G, K), gemmate (approximately spheroidal; Fig. 3Q), or verrucate (rounded and flattened, greater in diameter than height; Fig. 3R). Striate (parallel, elongated across the surface of the tectum) and rugulate (irregular, elongated across the surface of the tectum) supratectal elements were not observed in the present study. Supratectal elements can also be categorized by size (Character 18), which is here divided into two states, diameter greater than 1 μm (corresponding to macro-elements, usually visible under LM) and diameter less than 1 μm (micro-elements, usually only visible under electron microscopy).

Beneath the supratectal elements is usually a sculptured tectum. In LM studies, the sculpturing patterns are often categorized as OL- or LO-pattern following Erdtman (1952). Seen under the electron microscope, tectum sculpture forms are found to be numerous, varied, and not necessarily discrete or mutually exclusive; furthermore, their interpretation depends, to some extent, on the nature of both the supratectal elements above and the infratectum below (Punt et al., 2007). Here, we categorize states as follows (Character 19): perforate (with small, well-spaced openings less than 1 μm in diameter; Fig. 3L), foveolate (with large, circular, well-spaced holes greater than 1 μm diameter), fossulate (with elongate, irregular grooves; Fig. 3S), reticulate (with polygonal openings separated by narrow muri; Fig. 3T), areolate (with raised polygonal areas separated by narrow grooves; Fig. 3M), *Croton* L.-patterned (comprising groups of raised polygons arranged around a central space; Fig. 3U, V), rugulate (with irregular, elongate elements greater than 1 μm long; Fig. 3H), striate (with elongate parallel elements and grooves or spaces between; Fig. 3W), striato-reticulate (with elongate parallel elements and cross-links between), and imperforate (without sculpture, sometimes re-

ferred to as a "complete tectum" sensu Erdtman [1952]). There also exists a unique tectum sculpture reported only for *Amborella trichopoda* Baill. (treated in more detail by Lu et al., 2014, this issue), which comprises small cupules constructed of coiled cylindrical strands (Sampson, 2000).

Beneath the tectum lies the infratectum (Character 20), which is limited to one of three forms in almost all seed plant species (Van Campo & Lugardon, 1973): alveolate (restricted to gymnosperms and extinct, non-angiosperm seed plants), columellate (restricted to angiosperms; Fig. 3X), and granulate (observed in some species of angiosperms, most notably Annonaceae, and independently in some gymnosperms such as Gnetales). The infratectum itself may in some taxa (e.g., many Asterales) be distinguished into a number of layers (Character 21).

Below the infratectum are the foot layer (Character 22) and endexine (Fig. 3N), if present (Character 23). The endexine may differ in nature (Character 24: compact, spongy, or lamellar at maturity) and extent (Character 25: continuous, discontinuous, or solely apertural). In some taxa (e.g., Asteraceae), one or more exine layers are separated by a cavity known as a cavea, coded here as present or absent (Character 26). Also in some taxa (most notably Asteraceae), the elements that make up the exine are themselves perforated with tiny holes known as internal foramina (Character 27). In other taxa, the exine is traversed by minute, radially orientated channels (microchannels; Character 28). The latter two characters are usually visible only under TEM, and their presence or absence is rarely reported in the literature. In certain taxa (including many Asteraceae such as Fig. 3I), the exine may be folded into protruding ridges (lophae) and depressions (lacunae; Character 29), forming a pattern described as lophate (Wodehouse, 1935; Punt et al., 2007). Pollen that is not fully lophate but bears supratectal elements in a pattern as if upon lophae is described as sublophate.

External structures. Two types of structures external to the pollen grain, possibly of tapetal origin, have been highlighted in previous studies. Viscin threads (Character 30) are acetolysis-resistant, sporopollenin strands arising from the surface of the exine, common in some Onagraceae and Ericaceae. Orbicules (Character 31) are granules of sporopollenin found on the supratectal surface, sometimes referred to as Ubisch bodies. Although frequently reported from the surface of the anther locules, usually in species with a secretory tapetum, we record these only when they are found on the surface of the mature pollen grains themselves.

COMPARISON OF METHODS FOR STUDYING TRAIT EVOLUTION

In this paper, we test three aspects of the methodology for estimating ancestral states of pollen morphological characters upon a phylogeny of higher taxa: coding strategy, optimization method, and base tree topology.

Coding strategy. Because this study is based on a phylogeny of higher-level (ordinal) relationships in angiosperms inferred from character states (molecular sequence data) for individual species, we tested three methods of coding higher taxa for ancestral state optimization, represented in three different matrices: the species exemplar method (using mostly single exemplars; the possibility to select multiple exemplars was not available as the terminal taxa were preselected); the comprehensive method (where possible coding at least six species from across the phylogenetic and morphological range of each order, and displaying all observed states, with taxa selected to represent orders based on the delimitations of 2009); and the democratic method (sensu Bininda-Emonds et al. [1998]; a matrix created from the comprehensive matrix but with polymorphic data points reduced to the most common state. When two states were found to be equally prevalent, the state was coded as unknown). We did not test the ancestral method because neither fossils, ontogenetic data, nor existing phylogenies were widely available. We also created a modified species exemplar matrix in which polymorphic data points were reduced to the unknown state (?).

Coding strategies were compared for ancestral state optimization using the MP algorithm implemented in Mesquite versions 2.6–2.72 (Maddison & Maddison, 2009), because this is the only method of analysis currently applicable to matrices containing polymorphic cells, as found in two of the coding strategies tested. For consistency with tests of methodology, they were tested on a single phylogeny (Jansen et al., 2007). Consistency (ci) and retention indices (ri), as well as ensemble indices (CI, RI) were calculated for the four matrices on this phylogeny.

Optimization method. We tested four methods for optimization of ancestral character states: MP, ML, EB, and HB. Maximum parsimony optimization was conducted using Mesquite v.2.6–2.72 (Maddison & Maddison, 2009), with unordered (Fitch) parsimony under the "trace character history" option, and MacClade v.4.06 (Maddison & Maddison, 2000), using the "trace all possible changes" option. Maximum likelihood optimization was also conducted in Mesquite, applying an Mk-1 model, with param-

eters estimated from the data and using the default survey interval (two optimizations at a coarseness of 1.0 and 10.0, respectively). Empirical and hierarchical Bayesian optimizations were carried out using BayesMultistate (Pagel et al., 2004) in BayesTraits (<http://www.evolution.rdg.ac.uk/BayesTraits.html>). Results were examined in Microsoft Excel. Since the three model-based optimization methods are at present applicable only to datasets without polymorphic data points, to enable comparison across all four methods all were tested on the only two matrices that comply with this requirement: the democratic matrix and the modified species exemplar matrix (polymorphic data removed). The comprehensive matrix contained so many polymorphic cells that their removal would result in an unreasonably high proportion of missing data. For consistency with tests of coding strategy, they were tested on a single phylogeny (Jansen et al., 2007).

Testing multiple priors (see Appendix 4) indicated that the most appropriate methodology for HB inference with our data was to implement the reversible jump hyper prior (rjhp) mechanism as recommended in BayesTraits. This was applied for all characters, starting with a run of 100,000 iterations, sampling every 100 generations and a burn-in of 1000 generations. Rjhp settings (prior distribution type, range for uniform seeding distribution, and rate deviation) were manipulated for each character to ensure that the distribution of transition values included the point estimates obtained from the EB analysis, and so as to obtain a mean acceptance value of 20%–40% (minimizing autocorrelation among successive states of the chain while still exploring parameter space thoroughly; Pagel & Meade, 2006). Tests suggested that sampling every 300 generations successfully avoided autocorrelation between adjacent reported results. Plots of posterior values suggested that a burn-in period of anything greater than 10,000 generations was ample. Thus, the final parameter settings for the HB analysis were to sample every 300 generations, with a burn-in period of 20,000 generations for a net total of 5,000,000 generations (i.e., 5,020,000 including the burn-in period).

At each node, the favored states (i.e., the most parsimonious, most likely or those with the greatest posterior probability) generated by the four methods were compared. Where MP optimization gave multiple most parsimonious states, these were all considered equally supported and, thus, the method to be congruent with other methods if any of the states agreed. Similarly, where no single state was reconstructed for a node under ML (usually due to missing data), or where all states were found to be equally probable under Bayesian analysis, all states were considered to be possible and, therefore, the optimization to be congruent with any state generated by the other methods of analysis.

Starting tree topology. Five alternative arrangements were tested, representing the most common and significant variations found in recent, comprehensive phylogenetic studies (a wider range of topologies will be investigated in the second paper in this series (see Lu et al., 2015), focusing on basal angiosperms): (1) *Amborella trichopoda* sister to all other angiosperms, with the Nymphaeales the next branching clade, Chloranthales sister to magnoliids, and monocots sister to eudicots—this topology was obtained, with strong support, by Jansen et al. (2007; Fig. 2); (2) *A. trichopoda* sister to Nymphaeales, together sister to the remainder of angiosperms, Chloranthales, magnoliids and eudicots forming a polytomy and together sister to monocots—this topology was obtained by Soltis et al. (2007); (3) *A. trichopoda* sister to Nymphaeales (sensu Soltis et al., 2007), Chloranthales sister to magnoliids, and monocots sister to eudicots (sensu alternative topology in Jansen et al., 2007); (4) *A. trichopoda* sister to all other angiosperms (sensu Jansen et al., 2007), Chloranthales, magnoliids and eudicots forming a polytomy and together sister to monocots (sensu Soltis et al., 2007); (5) relative positions of both *A. trichopoda* and the Nymphaeales, and Chloranthales, magnoliids, monocots, and eudicots, left unresolved.

Topologies were compared for ancestral state optimization using the MP algorithm implemented in Mesquite versions 2.6–2.72 (Maddison & Maddison, 2009). Model-based methods of analysis were not appropriate because not all phylogenies were associated with branch lengths. For consistency with tests of methodology, this was conducted using the modified exemplar species matrix (polymorphic data removed).

Results

POLLEN CHARACTERS

After initial documentation of 35 palynological characters, three were removed due to extensive (greater than 50%) missing data (viz. aperture position in polar view, mesoaperture presence/absence, and endoaperture shape) and one due to invariability (tectum presence/absence). In the resulting matrix coded according to the exemplar method, five characters were parsimony-uninformative (and entirely invariant in the democratic matrix)

but were retained for their potential usefulness in model-based analyses (viz. Character 15, exine differentiation; Character 26, cavea; Character 27, internal foramina; Character 29, exine folding; Character 30, viscin threads). The final matrices (Appendix 5) contained 31 characters (Appendix 3). To enable complete comparisons across coding methods, all 31 characters were retained in the comprehensive and democratic matrices, even when they were invariant in these matrices.

COMPARISON OF METHODS FOR STUDYING TRAIT EVOLUTION

Coding strategy. The species exemplar matrix contained 137 cells (7%) missing data and 111 cells (6%) inapplicable, a total of 248 cells (13%) treated as missing data by the analysis programs. Seventy-six cells (4%) were polymorphic; when these cells were treated as missing data to facilitate comparisons across methods, the level of missing data increased to 326 cells (16%). The comprehensive matrix contained 42 cells (2%) missing data and 28 cells (1%) inapplicable, a total of 70 cells (4%) treated as missing data by the analysis programs. Eight hundred and twenty cells (41%) were polymorphic. The democratic matrix contained 97 cells (5%) missing data and 30 cells (2%) inapplicable, a total of 127 cells (6%) treated as missing data by the analysis programs. By definition, this matrix contained no polymorphic cells (see Table 1; Appendix 5). Although the ci varied between matrices, ri and rescaled consistency indices (rc) for characters followed generally similar patterns across the matrices, with the same characters having high or low indices in all three matrices (Table 2). Ensemble indices differed considerably, with CI 0.27, RI 0.5, and RC (ensemble rescaled consistency index) 0.14 for the exemplar species matrix, CI 0.51, RI 0.76, and RC 0.39 for the comprehensive matrix, and CI 0.35, RI 0.72, and RC 0.25 for the democratic matrix.

When analyzed using MP, the differences in inferred ancestral character state optimization between the three coding strategies were often large, both in terms of the location and number of character state changes inferred (Table 1). In general, the species exemplar matrix invoked the greatest number of character state changes, followed by the democratic matrix. For example, Character 2 (pollen grain polarity; Fig. 4) shows a complex pattern of multiple state changes when coded using exemplar species and optimized using MP (Fig. 4A). However, when coded comprehensively for orders, only a single state change is inferred, from heteropolar to isopolar (Fig. 4B). The exact position of the change within the phylogeny is ambiguous, occurring either at the root of the eudicots, or between Ranunculales and the rest of the eudicots. With democratic coding (Fig. 4C), the results are intermediate between these two extremes, with a clear transition from heteropolar to isopolar grains at the base of the eudicots, but further switches to apolar and isopolar grains seen closer to the tips of the phylogeny.

Optimization method. The two matrices used for comparison of optimization methods differed significantly. Of the 31 characters examined, 16 (52%) displayed different numbers of character states in the matrix generated by the democratic method of coding compared to that using exemplar species (Table 3), with the democratic matrix generally having fewer states than the exemplar species matrix. For eight characters, the democratic matrix contained only a single unambiguous state. In general, the democratic method of coding appeared to produce more consistent results across different methods of ancestral character state optimization: for the majority of characters (23), the democratic matrix generated fewer total disparities between methods than the species exemplar matrix (Table 4). The democratic matrix produced a greater number of differences for only two characters (Character 2, polarity, and Character 8, aperture number), while the remaining six characters generated the same total differences between methods with both matrices. Furthermore, the democratic matrix produced fewer differences at nodes for all six possible bilateral method comparisons. Therefore, the following comparison of methods is based on the species exemplar coding strategy, in order to avoid underestimating possible discrepancies between methods.

Notable differences were observed for many characters between the four different methods using the species exemplar matrix (Tables 3, 4; Fig. 5). By considering for each node the most parsimonious (from MP analysis), most likely (from ML analysis), and most probable ancestral state (from EB and HB analyses), it is possible to determine and quantify the differences between inferred ancestral states across all methods (Table 4). It should be noted, however, that such a comparison is relatively crude, taking no account of the relative likelihood (or probability) of each ancestral state, which in some cases was only marginally greater than that for the next most likely state. The level of congruence between optimization methods differed among characters. For example, four characters showed identical patterns with all methods tested: Characters 3 (symmetry), 13 (operculum), 21 (number of infratectum layers), and 29 (exine folding). Nine characters showed differences at more than 15 nodes (25%): Characters 5 (shape class), 6

Table 1. Comparison of pollen morphological character matrices examined in this study, generated using three alternative methods of coding higher taxa. C, comprehensive coding; D, democratic coding; E, species exemplar coding.

Character	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	Total
Number of states																																
C	5	3	2	2	5	5	5	7	3	2	5	2	2	4	2	2	4	2	11	3	2	2	2	3	3	2	2	2	2	2	2	n/a
D	1	3	2	2	2	4	3	5	3	2	2	2	2	3	2	2	4	2	7	2	2	2	2	3	3	2	2	2	1	2	2	n/a
E	3	3	2	2	4	4	5	7	3	2	4	2	2	4	2	2	3	2	11	3	2	2	2	3	3	2	2	2	2	2	2	n/a
Missing data (% cells)																																
C	0	0	0	0	6.3	0	0	0	0	0	0	1.6	0	0	0	0	14	14	1.6	1.6	1.6	1.6	3.1	23	11	1.6	14	17	0	0	0	3.6
D	0	0	1.6	6.3	6.3	3.1	4.7	1.6	3.1	0	4.7	4.7	4.7	4.7	0	4.7	19	17	14	1.6	7.8	1.6	4.7	31	16	1.6	14	17	0	0	3.1	6.4
E	0	0	1.6	0	13	0	0	0	4.7	4.7	4.7	20	4.7	4.7	0	0	50	50	3.1	1.6	4.7	22	22	48	45	1.6	36	45	0	0	0	12.5
Polymorphic cells (%)																																
C	69	58	45	16	70	89	94	92	39	48	81	53	45	63	0	83	64	56	95	6.3	6.3	16	20	7.8	9.4	6.3	7.8	0	6.3	3.1	3.1	41.3
D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E	0	1.6	0	6.3	7.8	7.8	20	13	1.6	0	4.7	4.7	1.6	3.1	0	4.7	3.1	1.6	27	1.6	0	0	3.1	0	0	0	0	0	0	0	6.3	3.8
Number of state changes inferred under MP criterion																																
C	0	2	2	3	2	5	1	2	2	1	0	4	0	2	1	2	3	2	5	2	0	1	0	7	2	0	0	2	0	0	0	53
D	0	6	5	5	4	11	8	5	5	5	4	4	1	6	1	10	8	4	9	1	0	2	0	6	2	0	0	2	0	0	0	114
E	3	12	5	5	8	13	18	15	11	7	13	9	4	13	1	15	9	5	24	2	1	3	4	8	4	1	1	2	1	1	2	220

Table 2. Consistency (ci), retention (ri), and rescaled consistency (rc) indices for pollen morphological characters examined in this study, for three alternative methods of coding higher taxa, calculated on the phylogeny provided by Jansen et al. (2007). Final columns show mean across all characters and ensemble indices (CI, RI, and RC) for each matrix. For the exemplar coded matrix, values given include polymorphic data points; values when polymorphic data points were treated as missing data were similar; ensemble indices were CI 0.27, RI 0.5, RC 0.14.

Character	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	Mean	Ensemble indices
Consistency index (ci)																																	
C	0	0.5	0.5	0.33	0.5	0.2	1	1	0.5	1	0	0.25	0	0.5	1	0.5	0.67	0.5	0.6	0.5	0	1	0	0.29	1	0	0	0.5	0	0	0	0.41	0.51
D	0	0.33	0.2	0.2	0.25	0.27	0.25	0.6	0.4	0.2	0.25	0.25	1	0.33	1	0.1	0.38	0.25	0.67	1	0	0.5	0	0.33	1	0	0	0.5	0	0	0	0.33	0.35
E	0.67	0.17	0.2	0.2	0.25	0.23	0.22	0.4	0.18	0.14	0.15	0.11	0.25	0.23	1	0.07	0.22	0.2	0.38	1	1	0.33	0.25	0.25	0.5	1	1	0.5	1	1	0.5	0.44	0.27
Retention index (ri)																																	
C	0	0.92	0.88	0.33	0	0.33	1	1	0.94	1	0	0.25	0	0.86	0	0	0.5	0	0.67	0.67	0	1	0	0.67	0	0	0	0.86	0	0	0	0.38	0.76
D	0	0.84	0.69	0.2	0	0.47	0.25	0.92	0.86	0.79	0.7	0.75	1	0.64	0	0.68	0.58	0.5	0.82	1	0	0.75	0	0.71	0	0	0	0.86	0	0	0	0.45	0.72
E	0	0.63	0.63	0.43	0.25	0.5	0.13	0.67	0.59	0.7	0.39	0.47	−.67	0.47	0	0.5	0.42	0.2	0.38	1	1	0	0.25	0.25	0	0	0	0.67	0	0	0.67	0.38	0.50
Rescaled consistency index (rc)																																	
C	0	0.46	0.44	0.11	0	0.07	1	1	0.47	1	0	0.06	0	0.43	0	0	0.33	0	0.4	0.33	0	1	0	0.19	0	0	0	0.43	0	0	0	0.25	0.39
D	0	0.28	0.14	0.04	0	0.13	0.06	0.55	0.35	0.16	0.18	0.19	1	0.21	0	0.07	0.22	0.13	0.55	1	0	0.38	0	0.24	0	0	0	0.43	0	0	0	0.2	0.25
E	0	0.1	0.14	0.09	0.06	0.12	0.03	0.27	0.11	0.1	0.6	0.05	0.17	0.11	0	0.03	0.09	0.04	0.14	1	1	0	0.06	0.06	0	0	0	0.33	0	0	0.33	0.14	0.14

Abbreviations: C, comprehensive coding; D, democratic coding; E, exemplar species coding.

(outline in polar view), 7 (size; Fig. 5A), 11 (ectoaperture shape), 12 (aperture membrane ornamentation), 14 (annulus; Fig. 5B), 17 (supratectal element shape), 19 (tectum sculpture), and 24 (endexine type; cf. Table 4). The greatest differences (> 35% of nodes) were seen for outline, size, annulus, supratectal element shape, and tectum sculpture. As an example, pollen size (Character 7; Fig. 5A) differed in optimization between methods at many key, deep nodes in the phylogeny. For instance, Node 6, which links *Pinus* L. and the angiosperms, had an inferred ancestral state of small under EB, medium under MP and ML, and large under HB. However, at more derived nodes (for instance, the root of the asterids and within this group), the four methods tended to converge on similar optimizations. In contrast, annulus type (Character 14; Fig. 5B) generated unanimous agreement between the methods at many of the deepest nodes in the phylogeny, such as the roots of angiosperms, monocots, and eudicots, all inferred as lacking an annulus.

In general, MP and ML tended to produce the most congruent results (differing at only 2.2% of possible nodes across all characters; Table 4); however, this may in part be due simply to the high number of equivocal optimizations generated by the ML analysis, which were treated as congruent with all possible states. The greatest incongruence was seen between the EB and the MP and ML methods (differing in 14.0% and 14.2% of optimizations, respectively).

Starting tree topology. For the five different topologies tested, using the species exemplar matrix (with polymorphic data points treated as missing data) and characters optimized with MP, very few changes were apparent in terms of ancestral state. The only characters for which any differences at all were noted were Characters 2 (polarity), 3 (symmetry), 4 (basic shape), 6 (outline in polar view), 9 (aperture position), 16 (supratectal elements), 19 (tectum sculpture), and 25 (endexine extent). In none of these instances did the differences affect key nodes (i.e., those deep in the phylogeny), and the number of state changes inferred was the same on all trees for all characters. The most significant difference noted was in Character 19 (tectum sculpture; Fig. 6). In this case, the state inferred at the root of the Nymphaeales differed depending on topology: for trees where *Amborella* Baill. and Nymphaeales were positioned on successive branches (e.g., Fig. 6A), the most parsimonious state for Nymphaeales was inferred as reticulate; for trees in which *Amborella* and Nymphaeales were sister groups (e.g., Fig. 6B), the most parsimonious state for the root of Nymphaeales was ambiguous, either reticulate or *Amborella*-type. This

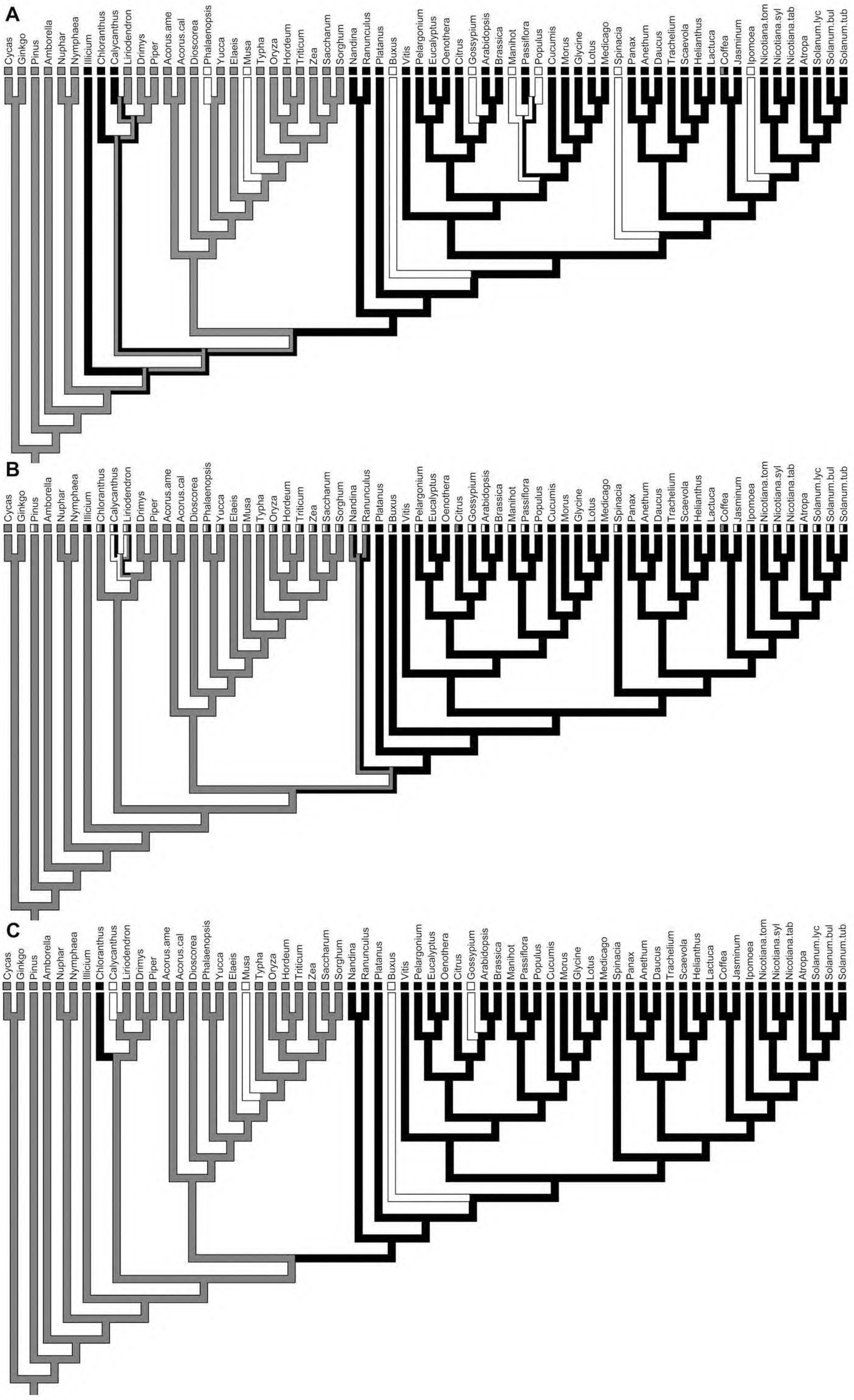


Figure 4. Comparing exemplar, comprehensive and democratic methods for coding higher taxa as terminals, using MP inference of ancestral states on angiosperm phylogeny, for the example of Character 2, pollen polarity: white, apolar; gray, heteropolar; black, isopolar/subisopolar. —A. Coding exemplar species as terminals (12 state changes inferred). —B. Coding taxonomic orders comprehensively as terminals (two state changes inferred). —C. Coding orders democratically as terminals (six state changes inferred). Branches displaying more than one shade indicate situations where the most parsimonious optimization is equivocal.

Table 3. Comparison of character statistics for matrices obtained using the democratic and exemplar (polymorphic cells treated as missing data—note that for this reason the number of states in the matrix may for some characters be fewer than the number of observed states) methods of coding higher taxa, calculated on the phylogeny provided by Jansen et al. (2007).

Character	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	Total
Number of unambiguous (i.e., neither inapplicable, '-', nor unknown, '?') states																																
D	1	3	2	2	2	4	3	4	3	2	2	2	2	3	2	2	4	2	7	2	1	2	1	3	3	1	1	2	1	1	1	n/a
E	3	3	2	2	3	4	5	7	3	2	3	2	2	4	2	2	3	2	10	3	2	2	2	3	3	2	2	2	2	2	2	n/a
Number of state changes inferred under MP criterion (figures in **bold** indicate greater than 100% difference in number of steps [MP analysis] between species exemplar matrix and democratic matrix. Figures in *italics* indicate where difference may be due to total number of character states rather than differences in optimization)																																
D	***0***	6	5	5	*4*	11	*8*	***5***	**5**	5	***4***	**4**	**1**	*6*	1	10	*8*	4	***9***	*1*	***0***	2	***0***	6	2	***0***	***0***	2	***0***	***0***	***0***	114
E	***3***	12	5	5	*8*	13	*14*	***15***	**11**	7	***13***	**9**	**4**	*12*	1	15	*9*	5	***21***	*2*	***1***	3	***4***	8	4	***1***	***1***	2	***11***	***1***	***2***	122
Number of internal nodes differing in inferred state between D and E under the four optimization methods (figures in **bold** indicate where more than a quarter of the 63 internal nodes show disagreement in reconstruction between the two coding strategies)																																
EB	11	4	0	3	**25**	**16**	**23**	**25**	2	11	**23**	13	1	**21**	0	**17**	**29**	8	**29**	3	2	6	11	4	1	3	3	0	3	4	8	309
HB	2	2	0	1	**22**	**16**	**26**	3	1	**17**	11	5	1	12	0	**16**	**28**	3	**27**	2	2	0	9	7	2	0	0	0	0	0	2	217
ML	0	0	0	0	1	9	3	0	0	**19**	1	3	1	6	0	10	**18**	2	8	1	0	0	5	1	0	0	0	0	0	0	0	88
MP	0	0	0	2	1	4	5	0	0	**21**	5	3	2	4	0	9	9	3	14	1	2	0	1	0	0	0	0	0	0	0	2	88
Total nodes affected	11	5	0	3	**33**	**28**	**29**	**25**	2	**25**	**24**	**16**	2	**26**	0	**19**	**35**	12	**39**	3	2	6	11	10	3	3	3	0	3	4	8	390
Total differences	13	6	0	6	49	45	57	28	3	**68**	40	24	5	43	0	52	**84**	16	**78**	7	6	6	26	12	3	3	3	0	3	4	14	704

Abbreviations: D, democratic matrix; E, species exemplar matrix; EB, empirical Bayesian inference; HB, hierarchical Bayesian inference; ML, maximum likelihood; MP, maximum parsimony.

Table 4. Comparison of observed node differences between optimization methods for matrices obtained using the democratic and exemplar (polymorphic cells treated as missing data) methods of coding higher taxa, calculated on the phylogeny provided by Jansen et al. (2007). Figures in **bold** indicate where more than 25% of the 63 internal nodes showed disagreement in reconstruction between the two coding strategies.

Character	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	Total
Number of observed differences (democratic matrix)																																
MP-ML	0	0	0	0	0	0	0	0	0	5	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	8
MP-EB	0	9	0	8	14	19	18	17	8	5	3	4	0	15	6	3	7	0	5	0	0	0	0	11	2	0	0	0	0	0	0	153
MP-HB	0	0	0	10	16	19	0	1	1	5	0	3	0	1	0	3	3	0	3	0	0	0	0	6	3	0	0	0	0	0	0	74
ML-EB	0	9	0	11	12	22	17	17	6	2	5	3	0	14	6	3	7	0	6	0	0	0	0	10	2	0	0	0	0	0	0	152
ML-HB	0	0	0	12	14	22	1	1	1	4	0	2	0	1	0	4	3	0	3	0	0	0	0	2	0	0	0	0	0	0	0	70
EB-HB	0	9	0	2	2	2	19	16	7	2	5	1	0	15	6	3	8	1	6	0	0	0	0	10	2	0	0	0	0	0	0	116
Total nodes affected	0	9	0	15	**16**	**23**	**19**	17	8	7	5	4	0	15	6	6	10	1	8	0	0	0	0	**17**	5	0	0	0	0	0	0	191
Total differences	0	27	0	43	58	84	55	52	23	23	13	14	0	46	18	17	27	1	23	0	0	0	0	40	9	0	0	0	0	0	0	573
Number of observed differences (exemplar species matrix)																																
MP-ML	0	0	0	7	0	1	1	0	0	0	3	2	0	0	0	6	13	1	8	0	0	0	0	0	0	0	0	0	0	0	0	42
MP-EB	11	7	0	7	21	28	26	6	8	13	16	13	0	20	6	6	19	10	10	2	0	6	4	11	3	3	3	3	0	4	6	272
MP-HB	2	0	0	0	2	14	20	2	2	5	7	5	0	5	0	6	17	4	17	1	0	0	2	5	0	0	0	0	0	0	0	116
ML-EB	11	9	0	14	18	25	24	8	8	13	16	13	0	22	6	11	16	7	15	2	0	5	4	11	3	3	2	3	0	4	5	278
ML-HB	2	2	0	13	3	12	19	4	2	5	9	6	0	8	0	11	5	2	11	1	0	0	2	3	0	0	0	0	0	0	0	120
EB-HB	9	7	0	1	19	14	9	10	6	8	12	9	0	17	6	2	22	6	17	1	0	6	2	12	3	3	3	3	0	4	6	217
Total nodes affected	11	9	0	14	**22**	**18**	**27**	11	8	13	**20**	**16**	0	**24**	6	13	**25**	10	**25**	2	0	6	4	**18**	3	3	3	3	0	4	6	324
Total differences	35	25	0	48	63	94	99	30	26	44	62	48	0	75	18	42	92	30	78	7	0	17	14	42	9	9	8	9	0	12	17	1053

situation is partly an artifact of removing polymorphic states: both Nymphaeales taxa were coded as missing data for this character in this matrix, due to both having polymorphic tectum sculpture types, which led to the most parsimonious state inferred at the root of the Nymphaeales depending entirely upon the surrounding taxa.

OVERVIEW OF ANGIOSPERM POLLEN EVOLUTION

Since topology was found to make relatively little difference to ancestral character state optimization at the level of angiosperms for the broad-scale pollen characters investigated, the following discussion is based on a single topology (Jansen et al., 2007), which is associated with meaningful branch lengths derived from molecular genetic state changes, enabling it to be used with both model-based and model-free methods. Although the method of coding higher taxa was found significantly to impact the results of ancestral character state optimizations (see above), for the purposes of exploring the evolution of pollen characters the following discussion will be based on the matrix produced using the species exemplar method (polymorphic characters treated as missing data); this avoids overestimation of congruence between methods.

A number of characters showed structured patterns when optimized on the phylogeny of angiosperms and, therefore, have potential to provide diagnostic or synapomorphic characters and hypotheses of evolutionary processes. Unfortunately, these tend also to be the characters that differed in optimization depending on the coding strategy, optimization method, or topology used. In general, characters that were unambiguously optimized in all contingencies were either those for which little data were available (e.g., Character 28, microchannels), or those known to be variable only within a single group, such as the cavea (Character 26), internal foramina (Character 27), and lophate exine (Character 29) of certain Asteraceae.

Nonetheless, of the 31 characters tested, only eight (Characters 5, 17, 18, 21, 25 to 27, and 29) were relatively uninformative. These were either variable only within a very small group (e.g., Character 29, exine folding), lacking in data (e.g., Character 17, supratectal element shape), or highly homoplastic at the present hierarchical level (e.g., Character 5, shape class). The remaining 23 characters were of some interest from an evolutionary or taxonomic point of view, despite variability across optimization methods. Several examples of particular interest are discussed below (Figs. 7, 8).

Character 2, pollen grain polarity (Fig. 7A), showed some differences in the ancestral character states inferred depending upon coding strategy, few differences among analysis methods, and no difference with tree topology. Optimization indicated a single inferred transition from heteropolar (the plesiomorphic state for angiosperms) to isopolar grains, most likely occurring on the internal branch between the stem of the monocots and the root of the eudicots (although it should be noted that further transitions are inferred on terminal branches, not shown in the figure, such as to isopolar in *Illicium* L., *Chloranthus* Sw., and *Calycanthus* L., and to apolar in *Ipomoea* L. and *Musa* L., among others).

Character 3, pollen grain symmetry (Fig. 7B), showed an entirely congruent pattern across all analysis methods and topologies tested, although optimizations differed with coding strategy. This character is inferred to have undergone two transitions from bilateral (the plesiomorphic state) to radial symmetry, once within monocots, on the branch leading to Zingiberales (represented by *Musa*) and Poales, and once on the internal branch between the stem of the monocots and the root of the eudicots. Again, there are further transitions within terminal branches (to radial symmetry in *Illicium*, *Chloranthus*, and *Drimys* J. R. Forst. & G. Forst.). However, radial symmetry seems to be fixed within eudicots, with no apparent reversals to bilateral symmetry among the taxa studied.

Character 4, basic shape (Fig. 7C), differed in reconstruction with tree topology and optimization method to a small extent but not with coding strategy. This character shows a clear transition from boat-shaped to globose occurring on the branch leading to Zingiberales and Poales, and another transition occurring some time after the divergence of the Nymphaeales from the rest of the angiosperms. With MP and ML, this latter change of state is inferred to have occurred before the divergence of the branch leading to Austrobaileyales, but with both Bayesian methods it is inferred prior to the root of the eudicots. Depending on the method, further transitions are invoked, e.g., to globose in the magnoliids, or a reversal to boat-shaped in early-diverging monocots. In all cases, globose pollen appears fixed within eudicots, with no apparent reversals to boat-shaped pollen. Due to high levels of missing data in basally branching taxa (polymorphisms within taxa removed for comparison of optimization methods), the plesiomorphic state at the base of the angiosperms was not resolved by all methods.

Character 6, outline in polar view (Fig. 7D), shows a much more complex pattern and displayed

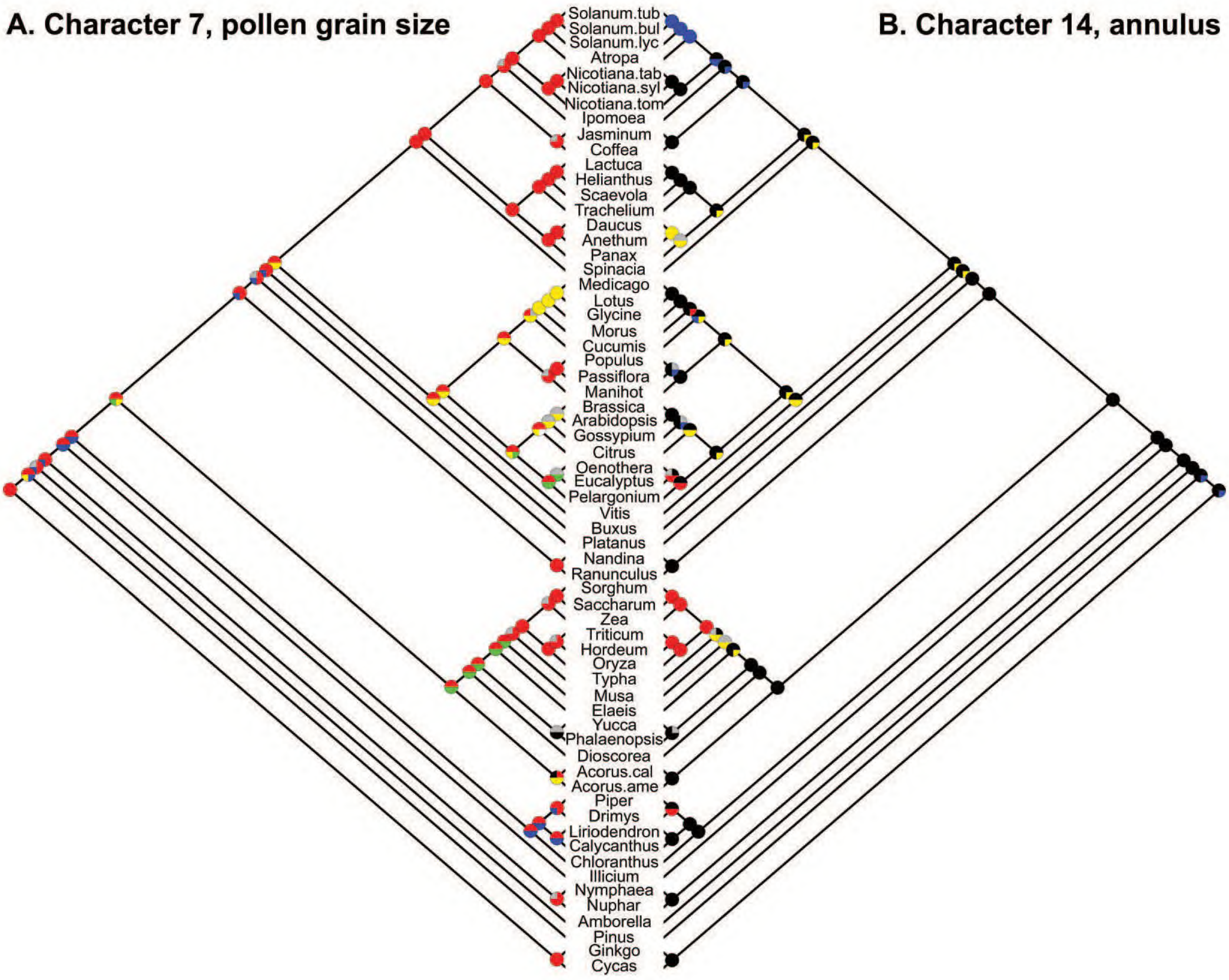


Figure 5. Comparing MP, ML, EB, and HB methods for inferring ancestral character states on angiosperm phylogeny, using species exemplar coding with polymorphic data points treated as missing data, for two example characters. Solid circles indicate congruence between all methods of optimization. Pie charts indicate incongruence: lower left segment, most parsimonious results; lower right segment, most likely result; upper right segment, most probable result of EB analysis; upper left segment, most probable result of HB analysis. —A. Character 7, pollen grain size: black, very small; yellow, small; red, medium; blue, large; green, very large. —B. Character 14, annulus: black, absent; yellow, costate; red, aspidate; blue, vestibulate. In both cases gray indicates an equivocal optimization.

differences with coding strategy, optimization method, and topology. The plesiomorphic state for angiosperms is resolved as oblate (although ambiguous in the ML optimization). In the two Bayesian optimizations, there is an inferred transition to circular outline before the branch leading to Austrobaileyales (represented by *Illicium*), in which case there is then a reversal to elliptic before the divergence of the monocots. All methods agree that there is a shift to circular grains within monocots and another before the root of the eudicots. However, in contrast to the previous characters, the situation within eudicots is highly complex and homoplastic, with little congruence between analysis methods on the exact placement of state transitions.

Character 8, aperture number (Fig. 8A), shows some discrepancies between optimization methods, particularly in inferring states at nodes toward the root of the tree. This character showed no differences in optimization between tree topologies but some differences between coding strategies. The plesiomorphic state for angiosperms (and several subsequent nodes) was inferred as mono-aperturate (MP and HB) or 4- to 6-aperturate (EB); the ML optimization was ambiguous at this point. All four methods were congruent in inferred state for the roots of monocots (mono-aperturate) and eudicots (tri-aperturate), with multiple transitions on terminal and subterminal branches to 4- to 6-, 7- to 12-, or poly-aperturate grains.

Character 9, aperture position (Fig. 8B), shows inferences differing slightly depending on coding strategy, optimization method, and tree topology. All methods inferred the plesiomorphic state for angiosperm apertures to be distal. Most methods (except EB) concurred on a likely transition to equatorial

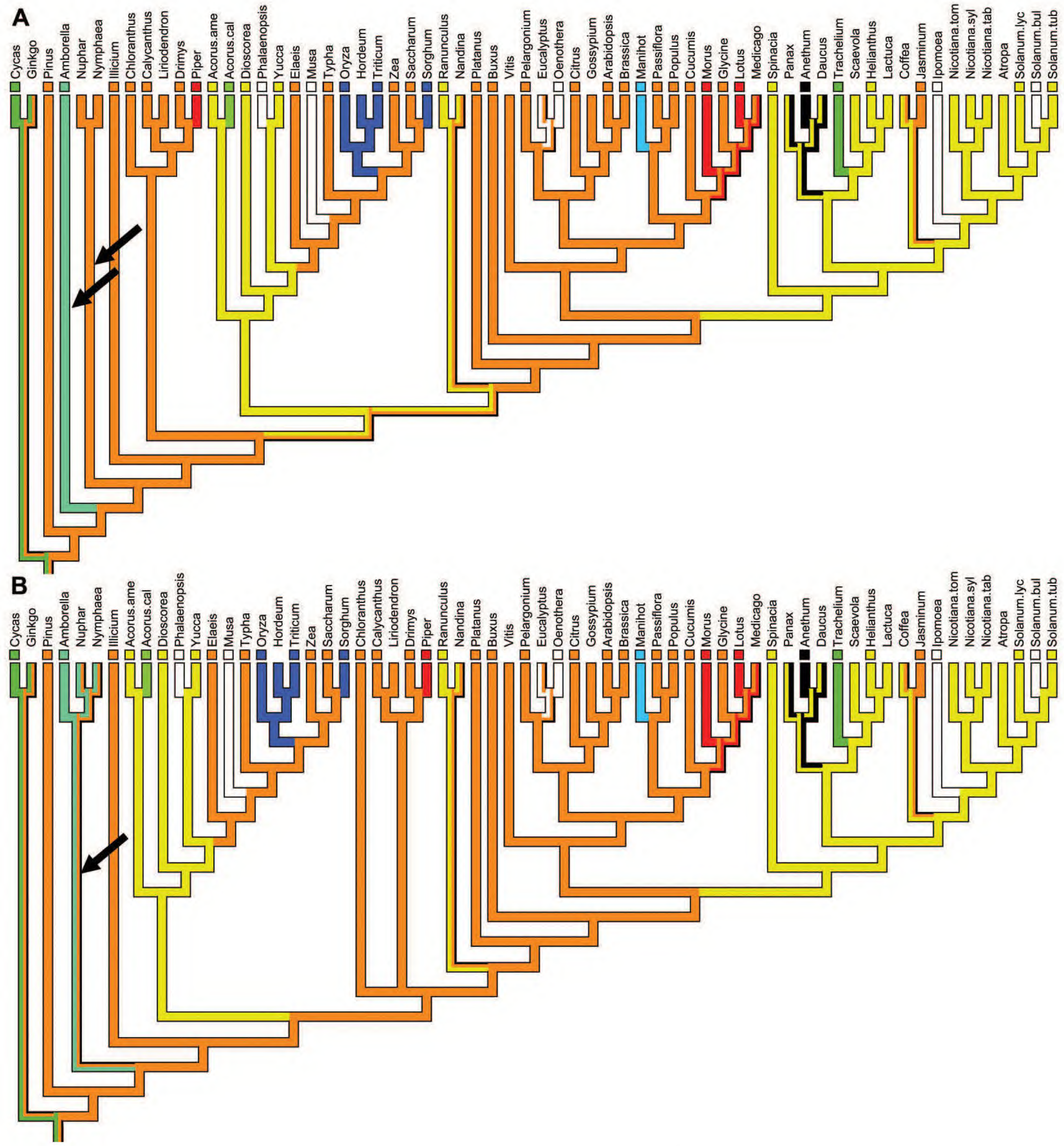


Figure 6. Comparing inferences of ancestral character state based on two different topologies for angiosperms, using the MP criterion and species exemplar coding with polymorphic data points treated as missing data, for the example of Character 19, tectum sculpture: white, imperforate tectum; dark blue, areolate; light blue, *Croton*-patterned; turquoise, *Amborella*-type; dark green, fossulate; light green, foveolate; yellow, perforate; orange, reticulate; red, rugulate; black, striate. Branches displaying more than one color indicate situations where most parsimonious optimization is equivocal. —A. Phylogeny as recovered by Jansen et al. (2007): *A. trichopoda* Baill. and Nymphaeales as successive sister groups to the remaining angiosperms. —B. Phylogeny as recovered by Soltis et al. (2007): *A. trichopoda* and Nymphaeales sister to one another. Arrows indicate groups where optimizations differ (see text for explanation).

apertures occurring on the branch leading to the root of eudicots. Multiple additional transitions to equatorial apertures are inferred on terminal taxa within the more basally branching angiosperms (such as *Illicium* and *Dioscorea* L.), and to global apertures within eudicots (e.g., *Gossypium* L., *Spinacia* L.). No transitions were inferred from distal to global apertures within the basally branching groups, nor any reversals to polar apertures within eudicots.

Optimization of Character 10, aperture structure (Fig. 8C), differed slightly between optimization methods, significantly with coding strategy, but not with tree topology. All methods concurred that the plesiomorphic state for angiosperms is simple

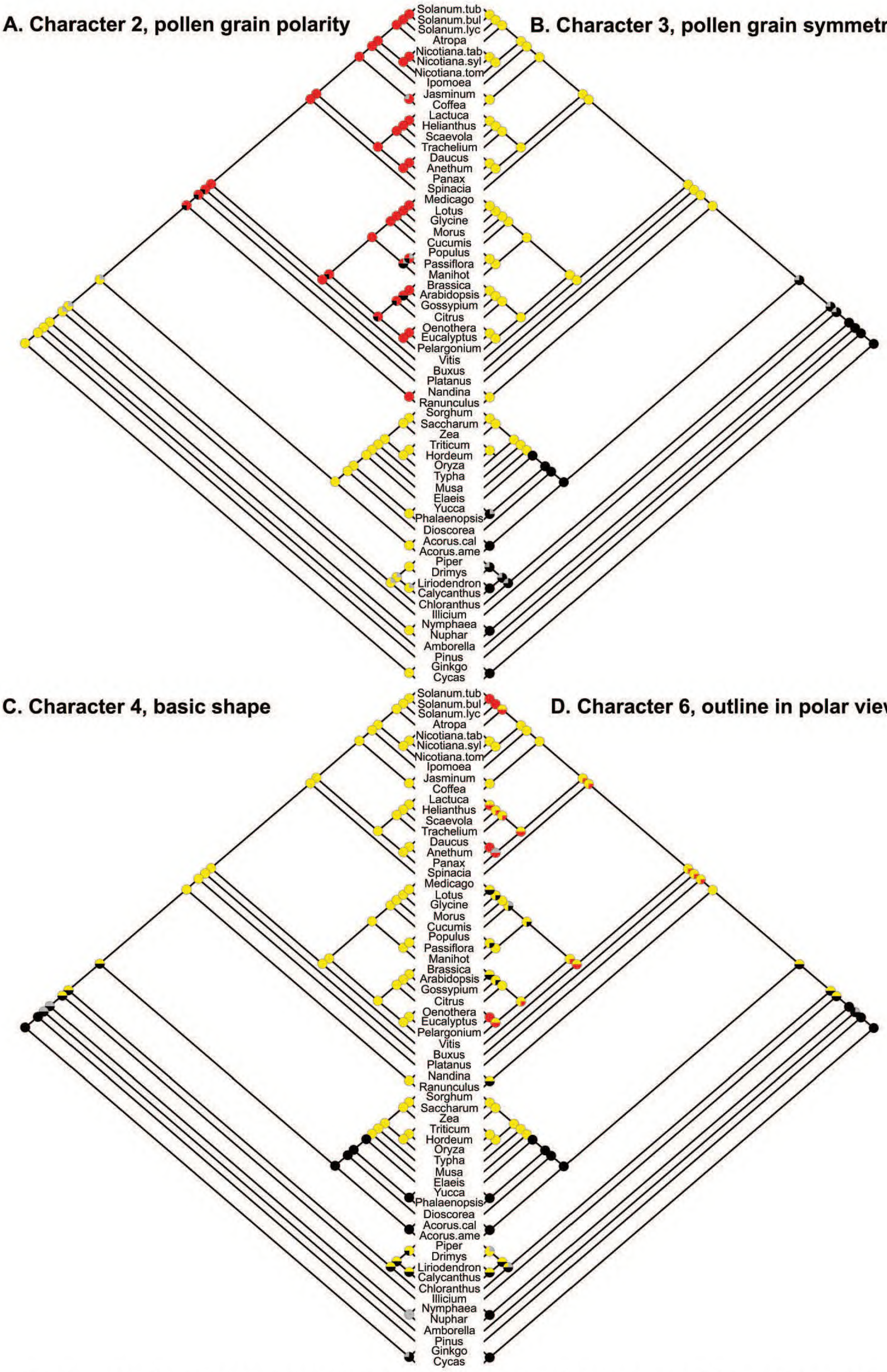


Figure 7. Examples of ancestral character states inferred on angiosperm phylogeny using all four analysis methods and the species exemplar coding method with polymorphic data points treated as missing data. Solid circles indicate congruence between all methods of optimization. Pie charts indicate incongruence: lower left segment, most parsimonious results; lower right segment, most likely result; upper right segment, most probable result of EB analysis; upper left segment, most probable result of HB analysis. —A. Character 2, pollen grain polarity: black, apolar; yellow, heteropolar; red, isopolar/subisopolar. —B. Character 3, pollen grain symmetry: black, bilateral; yellow, radial. —C. Character 4, basic shape: black, boat-shaped; yellow, globose. —D. Character 6, outline in polar view: black, elliptic; yellow, circular; red, polygonal; blue, lobate. In all cases gray indicates an equivocal optimization.

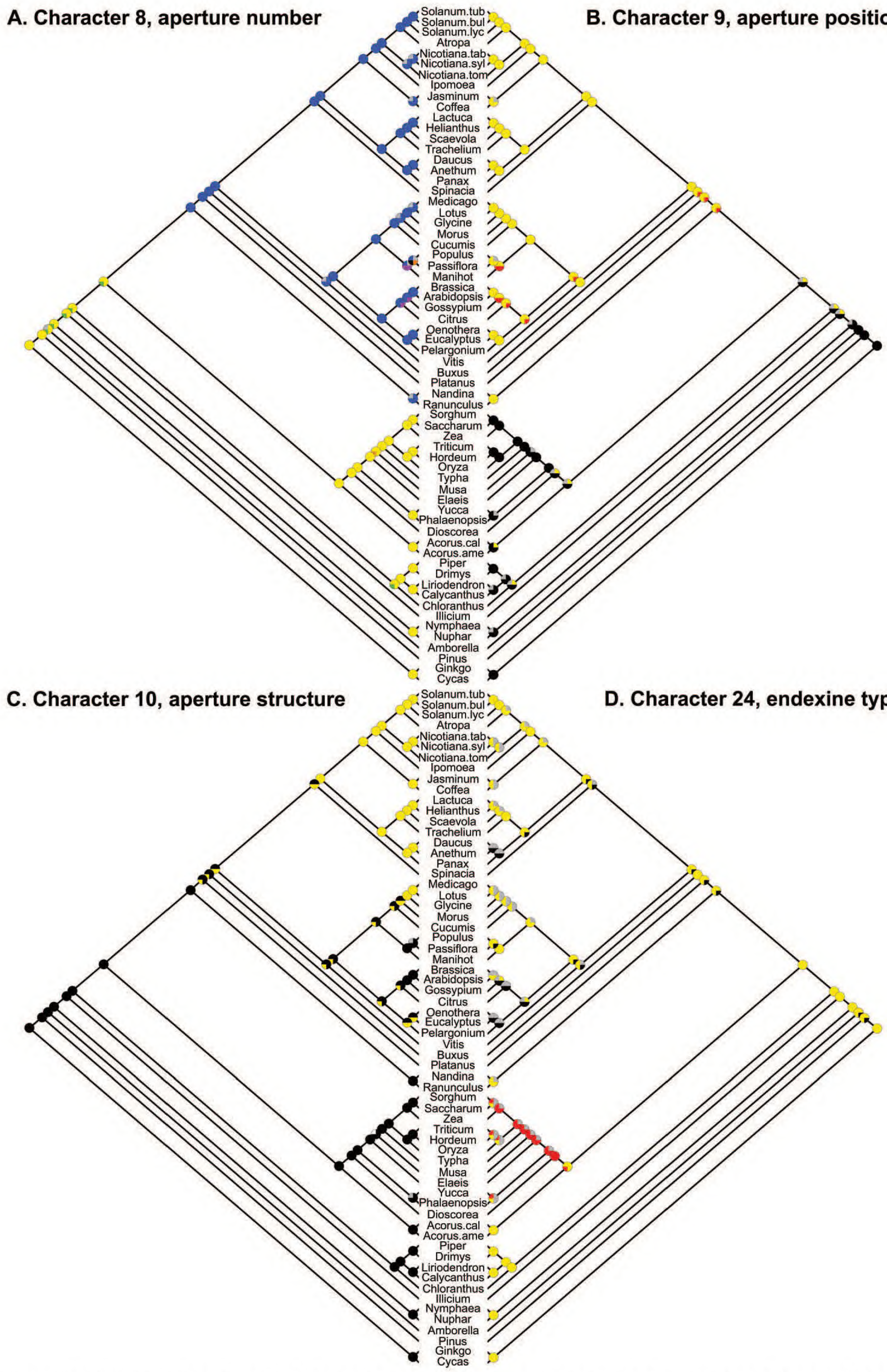


Figure 8. Further examples of ancestral character states inferred on angiosperm phylogeny using all four analysis methods and the species exemplar coding method with polymorphic data points treated as missing data. Solid circles indicate congruence between all methods of optimization. Pie charts indicate incongruence: lower left segment, most parsimonious results; lower right segment, most likely result; upper right segment, most probable result of EB analysis; upper left segment, most probable result of HB analysis. —A. Character 8, aperture number: black, zero; yellow, one; red, two; blue, three; green, four to six; orange, seven to 12; purple, many. —B. Character 9, aperture position: black, distal; yellow, equatorial; red, global. —C. Character 10, aperture structure: black, simple; yellow, compound. —D. Character 24, endexine type: black, compact; yellow, lamellar; red, granular/spongy. In all cases gray indicates an equivocal optimization.

apertures, a state that is retained until at least the separation of the basally branching eudicots (Ranunculales). However, the timing of the transition to compound apertures is more ambiguous. For example, under EB it occurs between the branch leading to Ranunculales and that leading to higher eudicots, under HB between Vitales (represented by *Vitis* L.) and higher eudicots, and under MP and ML only at the root of the asterids. Depending on the optimization method, further state changes are invoked on more terminal branches: either reversals to simple apertures (e.g., in *Oenothera* L. and *Spinacia*), or multiple origins of compound apertures (e.g., in Fabaceae and *Vitis*).

Our final example, Character 24, endexine type (Fig. 8D), was found to optimize in an entirely congruent manner across coding strategies and tree topologies, and to differ little between optimization methods. There are high levels of ambiguity in optimization due to the large amounts of missing data in this character; nonetheless, some interesting patterns are recorded. Most methods (with the exception of EB) concur that the plesiomorphic state for the angiosperm endexine comprises a lamellar structure at pollen grain maturity. Most methods also agree that there is a transition to granular/spongy endexine within monocots, above the branch leading to Acorales (with the exception of HB, which places this transition at the root of all monocots). The eudicots are inferred to be largely lamellar, with multiple transitions to a compact endexine structure.

When all characters are considered together in the context of the phylogeny, further patterns can be visualized. Figure 9 shows selected character state changes mapped using the MP (Fig. 9A) and HB (Fig. 9B) algorithms; ML and EB were not chosen for display due to their high rate of equivocal node reconstructions and large number of discrepancies with other methods, respectively. Maximum parsimony optimization inferred a total of 141 unambiguous character state changes (note that the vast majority of ambiguities resulted from missing terminal character states rather than multiple equally parsimonious optimizations). One hundred twenty of the inferred character state changes (85%) were autapomorphic at the family level or below and are excluded from Figure 9A for clarity. Hierarchical Bayesian optimization inferred a total of 283 unambiguous character state changes (with ambiguities due to missing data as well as to occasional occurrences of equal probabilities for each state at a node), 164 of which were (58%) autapomorphic at the family level or below, excluded from Figure 9B.

Figure 9A shows the 21 unambiguous character state changes that define interfamilial or higher-level groupings within angiosperms based on the MP optimization. Of the 45 nodes in the phylogeny (42 within angiosperms), 15 are associated with at least one inferred character state change, four are associated with two changes, and one with three changes (Fig. 9A). The node associated with the greatest number of inferred character state changes is that linking Zingiberales and Poales, which share transitions from bilateral to radial symmetry (Character 3), boat-shaped to globose pollen (Character 4), and elliptic to circular pollen in polar outline (Character 6), three characters which are not, of course, independent. The four nodes marked by two inferred character state changes are the root of the angiosperms itself, defined by a gain in supratectal elements (Character 16; albeit further modified many times within the phylogeny) and the transition from alveolate to columellate infratectum structure (Character 20); the root of the monocots, with transitions from globose to boat-shaped (Character 4) and circular to elliptic pollen (Character 6; both characters further reversed within the clade, in Zingiberales–Poales); the root of the commelinids (Arecales, Commelinales, Poales, and Zingiberales), characterized by transitions from a granular to smooth aperture membrane (Character 12) and perforate to reticulate tectum sculpture (Character 19; further modified several times on terminal branches within this clade); and the asterids + Caryophyllales (represented by *Spinacia*), defined by a further gain in supratectal elements (Character 16) and transition from reticulate to perforate tectum sculpture (Character 19). Key nodes characterized by a single inferred character state change under MP include that linking all angiosperms except *Amborella* and Nymphaeales (polar outline elliptic to circular, Character 6), that linking all angiosperms except the ANITA grade (Qiu et al., 1999; supratectal elements present to absent, Character 16), the root of the eudicots (aperture number one to three, Character 8), and the root of the asterids (aperture structure simple to compound, Character 10).

Figure 9B shows the 119 unambiguous character state changes that define interfamilial or higher-level groupings within angiosperms under HB optimization. Of the 45 nodes in the phylogeny, 36 are associated with at least one inferred character state change, five are associated with two changes, seven with three changes, five with four changes, two with five changes, and five with six changes, including *Phalaenopsis* Blume + *Yucca* L., which might represent the root of Asparagales, although sampling

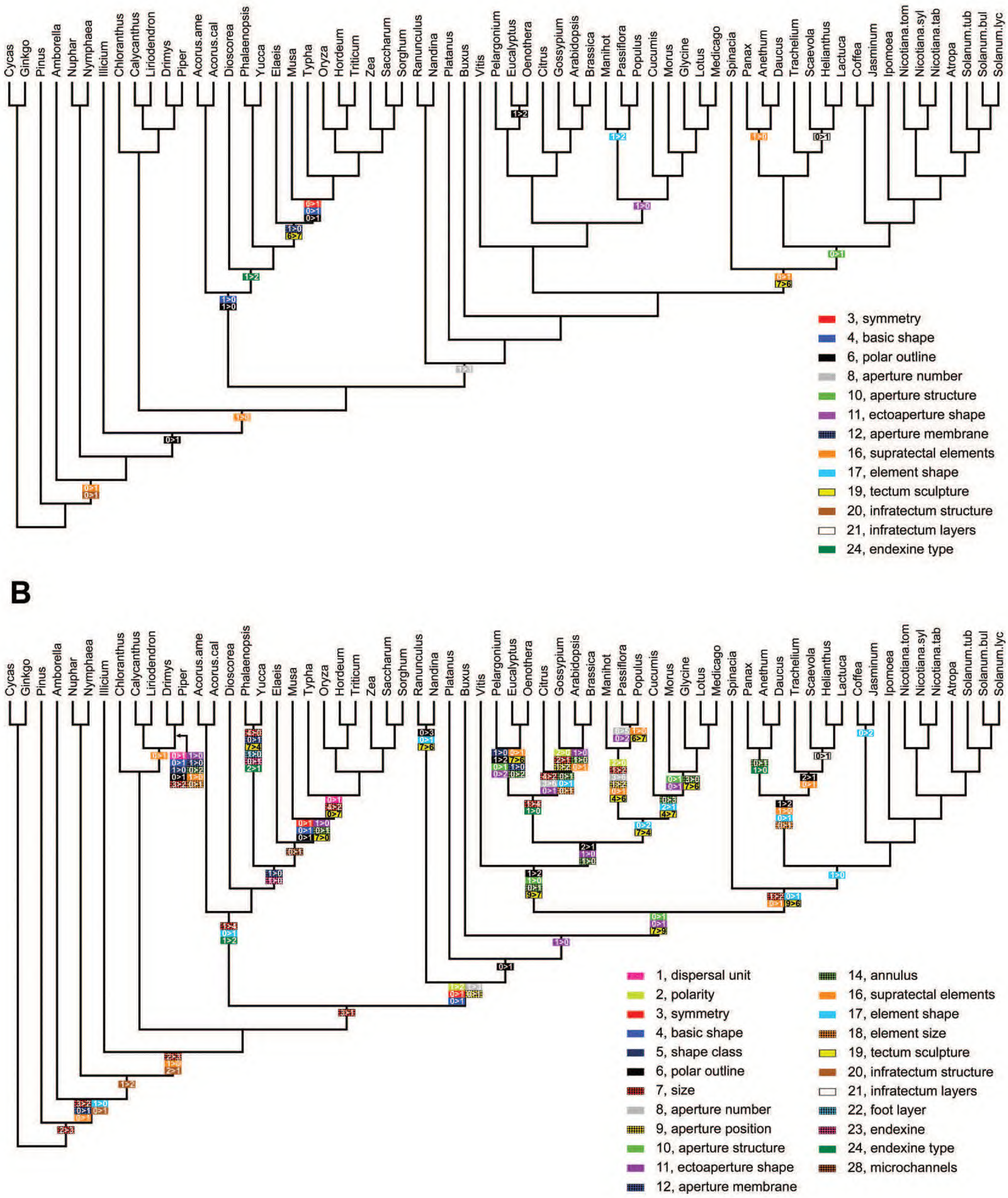


Figure 9. Summary evolution of pollen characters in angiosperms. Colors and patterns indicate characters (see legend for details), and numbers on bars indicate changes in state from left to right (for characters and states see Appendix 3). —A. The 21 unambiguous character state changes inferred using MP optimization of pollen characters at interfamilial level (changes autapomorphic for families or lower taxa removed), under species exemplar coding with polymorphic data points treated as missing data. —B. The 119 unambiguous character state changes inferred using HB analysis of pollen characters at interfamilial level (changes autapomorphic for families or lower taxa removed), under species exemplar coding with polymorphic data points treated as missing data. Changes are mapped assuming a single transition between differing states at adjacent nodes.

is minimal; and Zingiberales + Poales, characterized by porate, costate apertures (Characters 11, 14) and imperforate tectum sculpture (Character 19; although reversed at the next node), plus the three transitions observed in the MP analysis. One node is marked by eight character state changes (Geraniales [represented by *Pelargonium* L'Hér. ex Aiton] + Myrtales), and one with 10 changes (that linking Canellales [*Drimys*] and Piperales [*Piper* L.] within magnoliids). Two key nodes are characterized by five inferred character state changes: the root of the angiosperms, which in addition to the two transitions observed in the MP analysis is also marked by transitions from large- to medium-sized grains (Character 7), smooth to granulate aperture membranes (Character 12), and echinate to verrucate supratectal elements (Character 17)—all characters that change multiple times throughout the phylogeny; and the root of the eudicots, which, in addition to aperture number as observed in the MP analysis, is also distinguished by transitions from heteropolar to isopolar pollen (Character 2), bilateral to radial symmetry (Character 3), boat-shaped to globose (Character 4) grains, and distal to equatorial apertures (Character 9).

Discussion

In this first paper, we explore some of the available methods and issues to be considered when analyzing the evolution of morphological traits in a phylogenetic framework. Our analyses are based on a recent, comprehensive, model-based phylogeny for flowering plants (Jansen et al., 2007). Due to the patchy availability of genomic data, this phylogeny was heavily biased toward economically useful plants (such as Poaceae) and model taxa (such as *Arabidopsis* (DC.) Heynh. and *Nicotiana* L.). Nonetheless, it provided broad coverage of more than 50% of the orders and ordinal-level groups included in the APG III classification (APG III, 2009), and we believe this was a good starting point for study.

Comparison of methods for studying trait evolution

We tested three aspects of the methodology for inferring ancestral states upon a phylogenetic tree (coding strategy, optimization method, and tree topology), with 31 palynological characters. Certain characters, such as Characters 15 (exine differentiation), 28 (microchannels), and 29 (exine folding), were relatively robust to changes in the three parameters; these tended to be characters with low variability and high consistency indices, such as autapomorphies for one or a few taxa. Another set including Characters 6 (outline in polar view), 16 (supratectal elements), and 19 (tectum sculpture) showed high variability with all three parameters; these tended to be characters of high variability and homoplasy. However, many characters showed variability with some aspects of the methodology but not with others; these included Characters 2 (polarity), 7 (size), 12 (aperture membrane ornamentation), and 17 (supratectal element shape).

Coding strategy. The definition and coding of terminal taxa had a considerable impact upon ancestral character state optimization, due to its large resultant effect on the numbers of missing and polymorphic data points. Of the three coding strategies tested (species exemplar, comprehensive, and democratic), the greatest number of polymorphic data points was generated, as would be expected, by the comprehensive method, and the fewest by the democratic method (zero polymorphic data points as only one state was selected for each character in each taxon; cf. Table 1). The greatest amount of missing data occurred in the democratic matrix (due to removal of data points where states were of equal prevalence), and the lowest in the species exemplar matrix.

In general, the species exemplar matrix produced character state distributions with many more inferred changes of state than the other methods (Table 1; Fig. 4), while the comprehensive matrix resulted in the fewest inferred character state changes. This is clearly exemplified in Figure 4, for Character 2 (polarity), for which the species exemplar matrix generated 12 inferred changes of state, the democratic matrix six, and the comprehensive matrix two. This may be explained by the fact that the higher levels of polymorphism within terminal taxa in the comprehensive matrix are resolved, under MP, as changes within these terminal taxa, resulting in fewer character state changes on internal branches of the phylogeny (note that, for this reason, not all states shown in the matrices are always apparent in the figures; those that are autapomorphic for or within terminal taxa are not shown). By contrast, with the species exemplar coding strategy, any variability in the matrix occurs between terminal taxa, rather than within them, and thus is resolved as character state changes at internal nodes, increasing the number of inferred changes seen.

Characters that differed greatly when optimized based on different coding strategies tended to be major structural characters such as Characters 2 (polarity), 3 (symmetry), 10 (aperture structure), and 20 (infratectum structure). Aperture characters seemed to be particularly strongly affected, including Characters 10 (aperture structure), 11 (ectoaperture

shape), 13 (operculum), and 14 (annulus; cf. Table 1). A possible reason for this may be that aperture characters tend to be both highly informative at higher taxonomic levels and also variable at lower levels (i.e., within terminal taxa) resulting in differences in polymorphism levels with different coding strategies. We suggest that discrepancies due to coding strategy may be minimized in future studies with the use of probabilistic optimization methods, such as ML and Bayesian methods, which can accommodate proportional likelihoods at each node, and by increased taxon sampling, resulting in more accurate estimation of plesiomorphic states for higher-level taxa.

Optimization method. This study explores four methods (MP, ML, EB, and HB) of estimating character states for ancestral nodes and thereby inferring potentially diagnostic character states and evolutionary patterns. Previous studies have attempted empirically to compare the various techniques for ancestral state optimization in isolation, with mixed results. For example, Vanderpoorten and Goffinet (2006) found little difference between MP, ML, and Bayesian methods for morphological characters in the moss genus *Brachytheciastrum* Ignatov & Huttunen (Brachytheciaceae). In contrast, Ekman et al. (2008) found strong differences in the timing of the origin of the ascus in the Lecanorales (lichenized Ascomycota Caval.-Sm.), depending on the method used. Xiang and Thomas (2008) found differences between methods for highly homoplastic characters, but greater congruence for those with less homoplasy, when studying fruit and inflorescence characters in the Cornaceae. Wiens et al. (2007) found sometimes strongly supported and misleading differences between methods in a study of life history evolution in frogs.

Following an exploration of the impact of method under two coding strategies, we discuss the four methods of optimization in detail using only the species exemplar coding strategy. Our study found significant differences in ancestral state optimization between methods for most characters. Twenty-seven out of 31 characters differed to some extent, and 10 characters differed at more than 25% of internal nodes (Table 4). In general, the results of MP and ML optimizations were similar, with the results of the EB analysis differing most from the other three (Table 4). The MP method is perhaps most notable for the large number of ambiguous ancestral nodes recovered (Figs. 7, 8). The HB method produced the fewest ambiguous ancestral states, but it should be noted that states may be selected using this method, even if they have only a marginally greater inferred probability than alternative states (e.g., 51% compared to 49% for a binary character).

Figure 9B shows that the HB analysis resulted in many more character state changes that define groups (synapomorphies) than, e.g., MP (Fig. 9A); 42% of all character state changes inferred under HB define groups, compared to 15% under MP. This suggests that the HB method concentrates character state changes toward the base of the tree rather than on terminal branches (accelerates change), compared to MP. The EB method also tended to accelerate change compared to other methods (see, e.g., Fig. 8A, Character 8, in which aperture number increases at a much deeper node than with other methods).

The observed differences in ancestral character states reconstructed between optimization methods might be due to several factors, including optimality criteria, evolutionary models, and dependence (or independence) on branch lengths (in future studies using more than one base topology, this may also have an effect since only Bayesian methods can optimize character across more than one phylogeny at a time). Branch lengths seemed not to have a great effect: the number of character state changes inferred using ML and Bayesian methods was no greater on longer branches compared to MP. Most likely, the differences observed are due to the applied models of character state change, from the simple parsimony model, through the Mk-1 model used in ML and EB, to the complex multi-parameter continuous-time Markov models (Pagel & Meade, 2006) generated by HB. This raises the issue of whether the most complex models incorporating prior probabilities are based upon realistic assumptions, something of which some researchers are skeptical (e.g., Yang, 2006). Our tests applying a variety of priors (Appendix 4) indicated that, apart from the uniform, all priors produced generally similar results and, therefore, probably did not have a strong negative impact on any inferences of character evolution.

Figures 5, 7, and 8 demonstrate that in some cases the observed differences between methods would have a significant impact on inferences of character evolution or taxon diagnosis. For instance, Character 7 (size; Fig. 5A) was inferred under MP and ML to have evolved such that the ancestral condition toward the base of the tree was medium-sized pollen grains, while with Bayesian techniques many relatively basal nodes had large pollen size. For this character, the plesiomorphic states for angiosperms, monocots, and eudicots all differed between analysis methods (Fig. 5A).

Starting tree topology. For the phylogenies and characters explored here, tree topology had little

effect upon inferences of trait evolution using MP. Model-based optimization methods could not be used because branch lengths were unavailable for the multiple topologies, but we consider that the effect of topology would be similar regardless of the method used. Only eight characters (26%) showed slight differences in the manner in which their ancestral states were distributed (e.g., Fig. 6). In none of these instances did the differences affect key nodes (i.e., those toward the base of the tree), and the number of state changes inferred was the same on all trees for all characters. This is reassuring from the point of view of the analyses presented in this and future papers, because it indicates that, even where phylogenetic estimates may be uncertain, if the uncertainties involve only a few nodes and these are not within the regions of the phylogeny where character state changes are inferred to have taken place, there will be little negative impact upon inference of ancestral character states. Furthermore, where there are significant controversies in phylogeny estimation, and these do coincide with the likely timing of character state changes, multiple phylogenies with branch lengths in combination with probabilistic optimization methods can be used to explore all possibilities of trait evolution.

Since minimal differences were observed in the study presented here, the discussions that follow are based upon the single topology, from Jansen et al. (2007), for which branch lengths were available, making it possible to use both MP and model-based methods of optimization. We also focus on the exemplar species coding method, with polymorphic data points treated as missing data.

OVERVIEW OF ANGIOSPERM POLLEN EVOLUTION

Twenty-one (68%) of the characters examined showed relatively structured patterns when optimized on the phylogeny of angiosperms and, therefore, may have potential to provide diagnostic or synapomorphic character states, or to indicate key evolutionary transitions in pollen morphology. These included Characters 2 (polarity), 3 (symmetry), 4 (basic shape), 6 (outline in polar view), 8 (aperture number), 9 (aperture position), 10 (aperture structure), and 24 (endexine type; Figs. 7, 8). It should be noted that the many characters found to be highly homoplastic in our study (e.g., exine sculpture characters, size, shape in polar view) may also prove to be informative and diagnostic at lower levels with more detailed sampling.

Several characters show congruent patterns of inferred ancestral state distribution, with state changes located on the same branches of the phylogenetic tree. For instance, Characters 2 (polarity: apolar to heteropolar, with the exception of MP otimization; Fig. 7A), 3 (symmetry: bilateral to radial; Fig. 7B), 4 (basic shape: boat-shaped to globose, with the exceptions of MP and ML; Fig. 7C), 6 (outline in polar view: elliptic to circular, except MP and ML; Fig. 7D), 8 (aperture number: one to three, except MP; Fig. 8A), and 9 (aperture position: distal to equatorial, except ML and EB; Fig. 8B) all imply a change in state most likely on the branch subtending the root of the eudicots. These six character state changes that may be coincident with the origin of the eudicots are not independent but instead together form a switch in syndrome, from mono-aperturate (and, therefore, interpreted as heteropolar) to tri-aperturate grains in which the apertures are situated at the equator, rendering the pollen radially symmetrical, isopolar, and usually globose. A similar change in syndrome is suggested within monocots, where Poales and Zingiberales (represented by *Musa*) are linked by inferred changes from bilateral to radial symmetry (Character 3; Fig. 7B), boat-shaped to globose (Character 4; Fig. 7C) and elliptic to circular grains (Character 6; Fig. 7D), consistent wth a switch from boat-shaped grains with a single furrow to spheroidal grains with a single pore.

Key characters showing diagnostic changes at additional points on the phylogeny include aperture structure (Character 10) and endexine type (Character 24). For aperture structure (Fig. 8C), we can confidently infer a switch from simple to compound apertures within eudicots; however, the different optimization methods provide different conclusions on where this transition was most likely to have occurred: EB suggests it may have been as early as below the separation of the Proteales lineage from the remaining eudicots; HB suggests it occurred below the separation of the two large lineages corresponding approximately to rosids and asterids (with the caveat of limited taxon sampling); and MP and ML suggest that the main transition to compound apertures occurred just prior to the root of the asterids (with multiple additional occurrences in rosids). These differing interpretations provide a fruitful area for study because, while the adaptive function of compound apertures remains unclear, all possible optimizations suggest it could be linked to the evolution of highly diverse groups within angiosperms. The ambiguous optimization of this character may be partly dependent upon taxon sampling: e.g., the presence of known compound apertures within basally branching eudicot clades such as Buxales and Vitales implies that a single origin of compound apertures in the ancestor of all eudicots might be

most likely and would be in agreement with the interpretation of Blackmore and Crane (1998). This scenario adds yet another palynological synapomorphy for the eudicots. For endexine type (Character 24; Fig. 8D), there appear to be some clear patterns, including lamellar endexine in the more basally branching groups and a unique transition to granular or spongy endexine within monocots. In general, however, we are wary of drawing firm conclusions for characters such as this, with large proportions of missing, ambiguous, or polymorphic data, which may lead to artifactual results, such as the apparent occurrence of compact endexine at two of the deepest nodes when using EB analysis (Fig. 8D).

Aperture position (Fig. 8B) is an exceptionally interesting character because, despite the ambiguity and lability seen at certain nodes, some transitions were never observed to occur (within the sampling constraints of this study). For example, there are no occurrences within the outgroups, monocots, or basal angiosperms (either at terminals or internal nodes) of globally distributed apertures, and within the eudicots there are no observed reversals to distal apertures. This suggests that aperture position is a fundamental character, consistent with the fact that it is determined very early in pollen grain development by the nature of the meiotic tetrad.

In general, the palynological characters found to change state at key points in angiosperm evolution (particularly the origins of angiosperms and eudicots) tend to be those highlighted by previous authors (Wodehouse, 1935; Erdtman, 1952; Walker & Doyle, 1975; Nowicke & Skvarla, 1979). These include basic developmental characters such as polarity and symmetry, apertural characters, and the fundamental structure of the exine as granulate, alveolate, or columellate (Van Campo & Lugardon, 1973). However, additional characters also proved to show unexpected patterns. For example, in terms of shape, it was the simpler distinction between boat-shaped and rounded grains rather than quantitative shape classes (Fægri & Iversen, 1989), which was found best to distinguish groups.

Our results suggest that plesiomorphic pollen morphological states for the tree as a whole (i.e., in the common ancestor of the angiosperms and gymnosperms sampled) include monad-dispersed, heteropolar, bilaterally symmetrical, subspheroidal, medium-sized grains that are elliptic in polar view, with single, distal, simple, colpate apertures lacking an operculum or annulus. An exine is present, lacking supratectal elements, with a single-layered, alveolate infratectum; a foot layer and endexine are also present, the latter lamellated and continuous; and microchannels, viscin threads, and orbicules are absent. For a few characters the plesiomorphic state differed between optimization methods; e.g., MP recovered a plesiomorphic state of globose pollen, while HB suggested boat-shaped; HB implied that the plesiomorphic state for aperture membranes was smooth, while MP recovered it as granulate.

The highly variable pollen morphology seen in the most basally branching taxon of the ingroup, *Amborella* (polymorphic in terms of characters such as basic shape, outline, size, aperture number, ectoaperture shape, aperture membrane ornamentation, and endexine presence or absence), rendered recovery of plesiomorphic states for the angiosperms difficult, especially with conservative methods such as ML. Nevertheless, inferred plesiomorphic states for angiosperms themselves were found to include monad-dispersed, heteropolar, bilaterally symmetrical, subspheroidal, medium-sized grains that are elliptic in polar view, with single, distal, simple, colpate apertures having granulate membranes but lacking an operculum or annulus. An exine is present, with a reticulate tectum bearing echinate supratectal elements and a single-layered, columellate infratectum; a foot layer and endexine are also present, the latter lamellated and continuous; and microchannels, viscin threads, and orbicules are absent. Once again, MP recovered a plesiomorphic state of globose pollen, while HB suggested boat-shaped; MP implied that the plesiomorphic state for supratectal elements was echinate, while HB suggested verrucate.

From this we can see that angiosperms are distinguished from other seed plants by several synapomorphic pollen characters (Fig. 9), the least ambiguous being the presence of supratectal elements and columellate infratectum structure. Most of our analyses (except EB) confirm the discovery of Van Campo and Lugardon (1973) that columellate pollen defines and diagnoses the angiosperms. It is possible that the evolution of columellate structure, which incorporates a degree more complexity and flexibility than alveolate and granulate structures, is a key factor underlying the potential for the complex, multilayered exine structures seen in many groups of angiosperms (such as Asteraceae; Blackmore et al., 2009). This may in turn facilitate pollinator specificity, harmomegathy (the ability of pollen grains to change shape and size depending on hydration conditions), and other adaptive traits. It is important to note that the limited sample of taxa investigated here includes few angiosperms with granulate exine structure (key missing taxa include Annonaceae; Walker, 1971; Le Thomas & Lugardon, 1976), and

thus generates a simpler picture than might be revealed by fuller sampling. A deeper analysis of this particular character in a phylogenetic context is provided by Doyle (2009). None of the potential synapomorphies for angiosperms are ubiquitous; all display some reversals or further changes at lower hierarchical levels but nonetheless indicate a change in pollen type coincident with their origin.

Within angiosperms, there are several key points at which palynological characters may provide diagnoses or provoke evolutionary hypotheses (Fig. 9). The branch below the separation of Austrobaileyales (represented by *Illicium*) from the rest of eudicots is characterized by a switch from elliptic to circular grains (MP analysis; this state change is reversed at other levels in the phylogeny) or, under HB analysis, by two highly homoplastic character state changes (size large to medium, supratectal elements present to absent) and one very consistent character: infratectum structure granulate to columellate. The monocots also are supported by multiple characters: under MP they are characterized by boat-shaped grains with an elliptic polar outline (reversed in Zingiberales + Poales), under HB by very large size, echinate supratectal elements and granulate endexine. The grouping of Zingiberales + Poales is supported consistently across all methods by changes in symmetry, basic shape, and polar outline as previously described (Fig. 9) and, under HB optimization, three further characters: ectoaperture shape, annulus, and tectum sculpture. The other well-supported grouping is asterids + Caryophyllales, which are inferred to be characterized by supratectal elements and perforate tectum sculpture, as well as compound apertures and lack of an annulus in the HB optimization.

As discussed above, the clade perhaps best supported by pollen morphological synapomorphies is eudicots which share, at least under some methodologies, at least six linked character states (all reversed or modified in some terminal taxa). This agrees with, and expands upon, Donoghue and Doyle's (1989) tricolpate clade (note that there are of course many exceptions to tricolpate pollen as a diagnostic character, including tricolpate non-eudicots in Schisandraceae [Wang et al., 2009b] and *Nelumbo* Adans. [Banks et al., 2007], as well as eudicots with more or fewer apertures). Donoghue and Doyle did not specify whether they mean tricolpate with three meridional (equatorial) apertures sensu Erdtman (1943) or simply with three colpate apertures sensu Punt et al. (1994). In either case, based on our analyses, it is now possible to redefine the clade as isopolar, radially symmetrical, globose, and tri-zono-colpate. The independent evolution of tri-aperturate grains in Austrobaileyales (*Illicium oligandrum* Merr. & Chun), and also in Arecales (Harley & Dransfield, 2003) is confirmed by the fact that the three pores have a different developmental origin to those in most eudicots, being derived according to Garside's rule rather than Fischer's (Doyle et al., 1990). In fact, tri-aperturate grains derived following Garside's rule are observed to have evolved at least four times within angiosperms, in Austrobaileyales, Arecales, Proteales (Blackmore & Barnes, 1995), and Santalales (Maguire et al., 1974), as well as in the fossil pollen genus †*Aquilapollenites* Rouse, which may be associated with Santalales (Erdtman, 1971).

The eudicots comprise some 75% of all angiosperm species (Drinnan et al., 1994) and, given the several palynological synapomorphies now linked with the group, it is reasonable to ask whether pollen characters may in part be responsible for its diversity, perhaps by providing for more efficient production, transport, or fertilization of male gametes. It is also interesting to note that the origin, ca. 125 million years ago (Soltis et al., 2008), and subsequent diversification of eudicots was roughly coincident with that of certain groups of pollinating insects, such as bees, in the Early to Mid Cretaceous Period (e.g., Danforth et al., 2006).

FRAMEWORK FOR FUTURE DETAILED ANALYSES

To provide a framework for our future, detailed series of analyses of pollen evolution within angiosperms, we employed the most recent, comprehensive classification for the group (APG III, 2009), which comprises 70 orders, ordinal-level clades and unplaced taxa. To organize the orders and families of angiosperms as defined in this classification into rational groupings for analysis based on evolutionary relationship, we used the consensus tree published alongside this classification, which contains 63 orders and ordinal-level groups and was constructed from the results of several studies including Qiu et al. (1999), Graham and Olmstead (2000), Savolainen et al. (2000), Soltis et al. (2000), Bremer et al. (2002), Jansen et al. (2007), Moore et al. (2007), and Wang et al. (2009a). It is congruent with more recent studies such as those from Moore et al. (2010). This was modified (Fig. 10) to include four families—Icacinaceae, Oncothecaceae, Mettenusiaceae, and Vahliaceae—present in the APG III classification but not in the phylogeny presented alongside it (APG III, 2009). These are usually resolved in phylogenetic analyses to the lamiid clade (e.g., Soltis et al., 2007), although there is not strong support for any precise position.

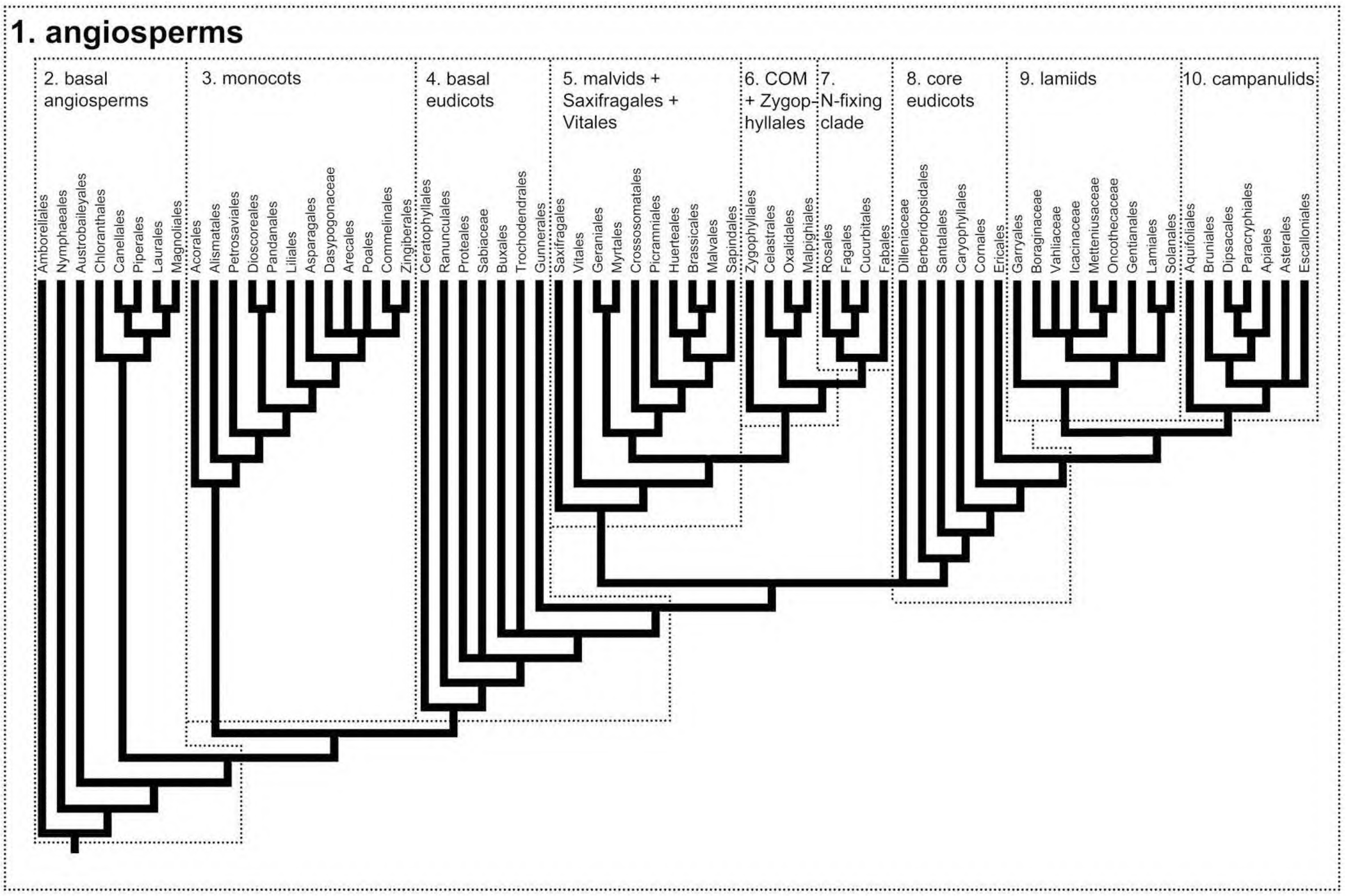


Figure 10. Framework phylogeny of angiosperm orders, divided into groups for subsequent analyses (for details see text). Formal and informal clade names taken from APG III (2009), Soltis et al. (1995), and Endress and Matthews (2006).

For the purposes of our study, they were placed with Boraginaceae as part of a polytomy at the base of the lamiids. This remains one of the most poorly resolved parts of the angiosperm tree, in need of more detailed phylogenetic analyses with larger amounts of data. Additional unplaced groups listed in the APG III classification (Apodanthaceae, Cynomoriaceae, *Gumillea* Ruiz & Pav., *Nicobariodendron* Vasudeva Rao & Chakrab., and *Petenaea* Lundell) were not included in our framework phylogeny since their likely affiliations were highly ambiguous; their pollen morphology will be analyzed in conjunction with their most likely relatives as discussed below.

We defined three criteria to divide the angiosperms into groups suitable for pollen evolutionary studies: (1) each group should be monophyletic or at least paraphyletic (never polyphyletic); (2) the groups should display similar ranges of pollen variability; and (3) they should be as similar as possible in size (number of species). Nine groups were defined (see Table 5; Fig. 10): basal angiosperms, monocots, basal eudicots + Ceratophyllales, malvids + Saxifragales + Vitales, COM (Celastrales, Oxalidales, Malpighiales) + Zygophyllaceae, the nitrogen-fixing clade, lamiids, campanulids, and other core eudicots.

A quick glance at Table 5 shows that these criteria were far from perfectly met but, most importantly, all groups provide a range of pollen variability that we consider feasible to study in a single paper, perhaps comparable to Nowicke and Skvarla's (1979) study of Centrospermae or Blackmore et al.'s (2009) on Compositae. Four of the nine groups defined are monophyletic (monocots, the nitrogen-fixing clade, lamiids, and campanuliids), with the remainder being paraphyletic. This is inevitable because, without sampling all orders individually, the removal of some monophyletic groups will always leave a paraphyletic subset. The criterion met least successfully was that of size, because this was considered subsidiary to the other two criteria. Thus, the size of the groups varies from ca. 6000 species (2% of angiosperms) in basal eudicots to 62,000 species (23%) in monocots. These size differences reflect the degree of pollen variation and depth of discussion required within the various groups. For instance, many monocots have relatively uniform pollen, while the pollen of basal eudicots is highly variable and likely to provide insights into the ancestral pollen type for both eudicots and core eudicots. To maintain consistent numbering across both groups and papers, including the present paper (Group 1, which deals with the angiosperms as a whole), the subsidiary groups are numbered 2 to 10 (Table 5; Fig. 10).

Group 2 (Table 5; Fig. 10; Lu et al., 2014), basal angiosperms, includes eight relatively small orders: Amborellales, Austrobaileyales, Canellales, Chloranthales, Laurales, Magnoliales, Nymphaeales, and Piperales, totaling 9335 species (3.5% of angiosperms). The group is paraphyletic, with a number of very long branches. This makes inference of ancestral states difficult and the use of model-based methods essential to allow multiple state changes on branches. With phylogenetic relationships still to some extent uncertain and implications for elucidating palynological traits underlying the diversification of the angiosperms as a whole, this is a crucial group.

Group 3 (Table 5; Fig. 10; Luo et al., in prep.), the monocots, is a monophyletic group comprising ca. 62,000 species (23% of angiosperms) in ca. 10 orders, including around 20,000 species in Poales and nearly 30,000 in Asparagales (22,500 within Orchidaceae). The relationships of most monocots are relatively clear, although the position and status of Dasypogonaceae are exceptions (APG III, 2009). The enigmatic genus *Ceratophyllum* L. (Ceratophyllaceae) is excluded from this group and placed in Group 4.

Group 4 (Table 5; Fig. 10; Zhang et al., 2015) contains the basal eudicots (Buxales, Gunnerales, Proteales, Ranunculales, Sabiaceae, and Trochodendrales) and *Ceratophyllum*, whose position as sister to eudicots is now relatively well supported (Jansen et al., 2007; Moore et al., 2007; APG III, 2009). Although small (ca. 6000 species, 2% of angiosperms), it is diverse in pollen morphology and may contain the answers to key questions underlying the explosive radiation of the eudicots. The positions of Buxales and Trochodendrales, and of Sabiaceae as sister to Proteales (and, therefore, its status as a separate order) remain uncertain (APG III, 2009).

Group 5 (Table 5; Fig. 10; Lu et al., 2015) is a fairly disparate, paraphyletic group of core eudicots, equating to the rosids excluding the monophyletic fabids (which are treated in Groups 6 and 7), with the addition of Saxifragales, which should perhaps be treated as a basal lineage of the rosids. Thus, Group 5 comprises Brassicales, Crossosomatales, Geraniales, Huerteales, Malvales, Myrtales, Picramniales, Sapindales, Saxifragales, and Vitales, a total of ca. 33,500 species (12.5% of angiosperms). Many of these lineages have only recently been defined and their relationships elucidated, e.g., Crossosomatales (APG III, 2009), Huerteales (Worberg et al., 2009), and Picramniales (Wang et al., 2009a). We also treat here the unplaced, isolated, monotypic genus *Petenaea*, identified in 1962 and limited to a small part of Central America, which is considered most likely a member of Malvales (Bayer et al., 1999; APG III, 2009), possibly close to Muntingiaceae. The pollen is reported as prolate, tricolporate, and microperforate (Bayer et al., 1999) but will merit further study to elucidate its relationships in more detail.

Groups 6 and 7 together make up the monophyletic fabids (Eurosid I). Group 6 (Table 5; Fig. 10; Luo et al., in prep.) comprises the monophyletic grouping of Celastrales, Malpighiales, and Oxalidales, collectively known as COM (Endress & Matthews, 2006; Zhu et al., 2007), plus Zygophyllales. The group contains ca. 20,000 species (ca. 7.5% of angiosperms). Here, we will also deal with two unplaced, isolated, monotypic genera, *Gumillea* and *Nicobariodendron*. The Peruvian species *G. auriculata* Ruiz. & Pav. was described in Cunoniaceae (Oxalidales) but is often considered a synonym of *Picramnia* Sw. (Picramniales; in Group 4) or placed in Simaroubaceae (Sapindales; in Group 5). It was maintained as separate by APG III (2009). Since it clearly belongs in one of the rosid groups, we will treat it here in Group 6 alongside Oxalidales. The pollen is little studied, although a very rough illustration is found in the work of Pavón (1798–1802: plate 245). *Nicobariodendron sleumeri* Vasudeva Rao & Chakrab. is possibly included in Celastraceae (Simmons, 2004). The species is endemic to the Nicobar Islands, very poorly collected, and the pollen has never been studied; it may provide very useful characters to help interpret the relationships of this enigmatic taxon but with such a limited distribution may be difficult to obtain.

Group 7 (Table 5; Fig. 10; He et al., in prep.), known as the nitrogen-fixing clade (Soltis et al., 1995, 1999), comprises Cucurbitales, Fabales, Fagales, and Rosales, together a monophyletic and strongly supported group in the recent analysis of Wang et al. (2009a). It contains ca. 31,000 species (11.5% of angiosperms). Here, we will also treat Apodanthaceae, a holoparasitic family of three genera, *Apodanthes* Poit., *Berlinianche* Vattimo, and *Pilostyles* Guill. (Mabberley, 2008). Apodanthaceae was previously included in Rafflesiaceae, from which it is now excluded (Blarer et al., 2000) and has since remained unplaced. Molecular data suggest a position either in Malvales or Cucurbitales (Nickrent et al., 2004). Cucurbitales now seems the more likely placement (Barkman et al., 2007; Filipowicz & Renner, 2010); thus, Apodanthaceae will be treated with Group 7. The pollen morphology of Apodanthaceae has been well studied (Takhtajan et al., 1985; Roubik & Moreno, 1991; Blarer et al., 2004) and, although highly variable within the family (ranging from inaperturate to tricolpate; Blarer et al., 2004), confirms its exclusion from Rafflesiaceae. However,

Table 5. Division of angiosperms (Group 1) into nine groups for subsequent palynological studies.

Group	Orders and ordinal-level clades contained	Approximate number of species	Approximate % angiosperms
2. Basal angiosperms	Amborellales, Austrobaileyales, Canellales, Chloranthales, Laurales, Magnoliales, Nymphaeales, Piperales	9335	3.5
3. Monocots	Acorales, Alismatales, Arecales, Asparagales, Commelinales, Dasypogonaceae, Dioscoreales, Liliales, Pandanales, Petrosaviales, Poales, Zingiberales	62,000	23
4. Basal eudicots + Ceratophyllales	Buxales, Ceratophyllales, Gunnerales, Proteales, Ranunculales, Sabiaceae, Trochodendrales	6000	2
5. Malvids + Saxifragales + Vitales	Brassicales, Crossosomatales, Geraniales, Huerteales, Malvales, Myrtales, Picramniales, Sapindales, Saxifragales, Vitales	33,500	12.5
6. COM + Zygophyllales	Celastrales, Malpighiales, Oxalidales (COM; Endress & Matthews, 2006; Zhu et al., 2007), Zygophyllales	20,000	7.5
7. Nitrogen-fixing clade	Cucurbitales, Fabales, Fagales, Rosales	31,000	11.5
8. Other core eudicots	Berberidopsidales, Caryophyllales, Cornales, Dilleniaceae, Ericales, Santalales	25,500	9.5
9. Lamiids	Boraginaceae, Garryales, Gentianales, Icacinaceae, Lamiales, Metteniusaceae, Oncothecaceae, Solanales, Vahliaceae	47,000	17.5
10. Campanuliids	Apiales, Aquifoliales, Asterales, Bruniales, Dipsacales, Escalloniales, Paracryphiales	33,000	12.5

pollen characters of Apodanthaceae have not been considered in the context of related families, and it is hoped further study might improve our understanding of the relationships of this family.

Group 8 (Table 5; Fig. 10; Wang et al., in prep.) comprises a paraphyletic grade of core eudicots that subtend the sister grouping of lamiids and campanulids. Together, the Berberidopsidales, Caryophyllales, Cornales, Dilleniaceae, Ericales, and Santalales include ca. 25,500 species (9.5% of angiosperms) with some interesting palynological features (such as pollen dispersed in tetrads in many Ericaceae). Cynomoriaceae, a holoparasitic, monogeneric family of one (Mabberley, 2008) or two species (Nickrent et al., 2005) from eastern Asia and the Mediterranean, has been placed variously in Saxifragales (Nickrent et al., 2005) or Santalales (Jian et al., 2008), sometimes under Balanophoraceae (e.g., Cronquist, 1981). Saxifragales and Santalales are quite separate in recent classifications (APG III, 2009) and fall into two different groups in our framework. Here, we will treat Cynomoriaceae in Group 8 alongside Santalales, as suggested (albeit poorly supported) by the most recent molecular results (Jian et al., 2008). Pollen of *Cynomorium* L. has not been studied since 1975, and previous studies have focused on pollen development rather than pollen grain morphology (Steindl, 1945; Terekhin et al., 1975), so further studies would be highly beneficial.

Groups 9 (lamiids) and 10 (campanulids) are relatively well-supported, monophyletic sister groups. The lamiids (Group 9; Table 5; Fig. 10; Wortley et al., in prep.) comprise ca. 47,000 species (17.5% of angiosperms), including Boraginaceae, Garryales, Gentianales, Icacinaceae, Lamiales, Metteniusaceae, Oncothecaceae, Solanales, and Vahliaceae. Pollen morphology within this group is highly variable, e.g., in Acanthaceae (Lamiales; Scotland, 1992). A number of families within this group (Boraginaceae, Icacinaceae, Metteniusaceae, Oncothecaceae, and Vahliaceae) remain of uncertain position and status (APG III, 2009). Again, it is hoped that pollen morphological study may shed light on their relationships.

The campanulids (Group 10; Table 5; Fig. 10; La et al., in prep.) include Apiales, Aquifoliales, Asterales, Bruniales, Dipsacales, Escalloniales, and Paracryphiales, a total of ca. 33,000 species (12.5% of angiosperms). Within the group, the positions of Escalloniales and Bruniales remain somewhat uncertain (APG III, 2009). More than half the group is included in a single family, Asteraceae (ca. 23,600 species), a family that has already been subject to detailed study of pollen morphological characters, which were found to be very useful for defining and diagnosing clades within it (Blackmore et al., 2009).

The methodology for treating each of these nine groups will be based on the results of the present study. Thus, base phylogenies will provide broad coverage of all orders and most families and will be based on molecular genetic data, ideally multiple loci, analyzed using model-based methods and presented with branch lengths. If, as in the present study, any unstable nodes are confined to morphologically homogeneous regions of the tree, a single phylogenetic estimate may be used as a base. However, in cases where there are significant controversies between topologies, it will be prudent to test multiple phylogenies. For paraphyletic groups, the overview phylogeny will contain representatives of the segregate taxa that render the study group paraphyletic, so that the phylogeny as a whole is monophyletic for the purposes of the study. For groups where a suitable overview phylogeny cannot be found, the appropriate section of a broader study such as the multi-gene analyses from Soltis et al. (2007, 2011) may be used; these include representatives of all orders in APG III (2009). Although it might be possible to build supertrees from more densely sampled smaller trees, Torices (2010) has shown that errors in branch length reconstruction on supertrees can lead to differences in the inference of character states, at least with ML. Additional, more comprehensive phylogenies may be selected for the constituent orders, sampling at least to family and possibly to generic level depending on the size and variability of the group, and coding for more specialized characters, the nature of which will depend on the palynological makeup of each particular group.

We have shown that coding strategy can have a significant effect upon ancestral character state reconstruction, perhaps most significant in characters that are informative at higher taxonomic levels but also variable at lower levels (i.e., within terminal taxa). For future papers, we recommend the use of democratic coding, which provides a good compromise between levels of polymorphism and missing data. However, the impacts of coding, as seen in this study, could be minimized by dense sampling and model-based optimization methods; if for this reason coding strategy is no longer considered to be an issue, the simpler species exemplar method may be used. Because optimization method can also have a significant effect upon ancestral character state reconstruction, probably due to differing models of evolution, future studies should employ at least one method with a simple optimization model (MP or ML) and one with a complex model (HB) for comparison.

Conclusions

The APG classification (APG III, 2009) is fast becoming the baseline of choice for robust evolutionary investigations of all kinds (for a recent example, see Hawkins et al., 2011). Our APG-based preliminary analysis of pollen characters in angiosperms immediately suggests diagnostic and synapomorphic characters for key monophyletic groups including angiosperms themselves and eudicots, as well as potential hypotheses of evolutionary process. In combination with recent successful family-level studies (e.g., Blackmore et al., 2009), it suggests that further, detailed analyses should be a priority for palynological research, to facilitate diagnosis of groups, indicate potential relationships for unplaced or unstable taxa, and uncover correlations with evolutionary/ecological events. The identification of key ancestral state combinations may also facilitate more accurate placement of fossil pollen grains in a phylogenetic context, supporting the dating and calibration of phylogenetic trees. Figure 10 suggests a number of taxa for which pollen morphology will be of particular interest in determining evolutionary patterns, such as many basal angiosperms, Acorales (the first branching clade of monocots), Ceratophyllales (sister to eudicots), Ranunculales (the first branching lineage of eudicots), Trochodendrales (the potential sister group to core eudicots), Gunnerales (the first branching lineage in core eudicots), Saxifragales (sister to rosids), Vitales (the first branching lineage in rosids), and many of the taxa in Group 8 which are potential sister groups to asterids or larger clades. Some of these, including Acorales (Wang et al., 1995) and Gunnerales (Wanntorp et al., 2004), have been relatively little studied.

Our work provides a reference for future palynological and systematic studies, a stimulus for angiosperm-wide analysis of other morphological character sets, and a methodological protocol for such studies to proceed. In terms of the latter, this paper shows that the results of ancestral trait optimization can be unstable when variables such as topology, coding, and evolutionary model are changed. Tree topology had little effect upon ancestral character state inference where discrepancies involved only nodes without inferences of character state change. Coding strategy can have a significant effect upon ancestral character state inference, probably due to differing levels of missing data and polymorphism between coding strategies. This effect is perhaps most significant in those characters that are informative at higher taxonomic levels but also variable at lower levels (i.e., within

terminal taxa). For such characters, species exemplar coding results in a greater number of inferred character state changes on internal branches within the phylogeny, while comprehensive coding results in fewer such changes. Finally, optimization method can also have a significant effect upon ancestral character state optimization. Until it becomes clear which models and priors are most appropriate, comparison of multiple methods with different assumptions is the most prudent approach to robust optimization.

Literature Cited

Acosta Castellanos, S. & A. E. Vilela. 1998. Anatomía folia y morfología del polen de *Drimys granadensis* var. *mexicana* (Winteraceae: Magnoliales). Polibotanica 8: 1–12.

Al-Wadi, H. M. & G. M. A. Lashin. 2007. Palynological and cytological characters of three species of genus *Solanum* (family: Solanaceae) from Saudi Arabia. J. Biol. Sci. 7: 626–631.

Alvarez, A. & E. Köhler. 1987. Pollen morphology of Agavaceae and some related genera. Grana 26: 25–46.

Amela García, M. T., B. G. Galati & A. M. Anton. 2002. Microsporogenesis, microgametogenesis and pollen morphology of *Passiflora* spp. (Passifloraceae). Bot. J. Linn. Soc. 139: 383–394.

APG. 1998. An ordinal classification for the families of flowering plants. Ann. Missouri Bot. Gard. 85: 531–553.

APG II. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Bot. J. Linn. Soc. 141: 399–436.

APG III. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Bot. J. Linn. Soc. 161: 105–121.

Audran, J.-C. & E. Masure. 1978. La sculpture et l'infrastructure du sporoderme de *Ginkgo biloba* comparées à celles des enveloppes polliniques des Cycadales. Rev. Palaeobot. Palynol. 26: 363–387.

Banks, H. I., P. J. Stafford & P. R. Crane. 2007. Aperture variation in the pollen of *Nelumbo* (Nelumbonaceae). Grana 46: 157–163.

Barkman, T. J., J. R. McNeal, S.-H. Lim, G. Coat, H. B. Croom, N. D. Young & C. W. De Pamphilis. 2007. Mitochondrial DNA suggests at least 11 origins of parasitism in angiosperms and reveals genomic chimerism in parasitic plants. BMC Evol. Biol. 7: 248.

Bayer, C., M. F. Fay, A. Y. De Bruijn, V. Savolainen, C. M. Morton, K. Kubitzki, W. S. Alverson & M. W. Chase. 1999. Support for an expanded family concept of Malvaceae within a recircumscribed order Malvales: A combined analysis of plastid *atp*B and *rbc*L DNA sequences. Bot. J. Linn. Soc. 129: 267–303.

Beasley, C. A. 1975. Developmental morphology of cotton flowers and seed as seen with the scanning electron microscope. Amer. J. Bot. 62: 584–592.

Behnke, H. D. 1997. Sarcobataceae. A new family of Caryophyllales. Taxon 46: 495–507.

Bininda-Emonds, O. R. P., H. N. Bryant & A. P. Russell. 1998. Supraspecific taxa as terminals in cladistic analysis: Implicit assumptions of monophyly and a comparison of methods. Biol. J. Linn. Soc. 64: 101–133.

Blackmore, S. 1982. Palynology of subtribe Scorzonerinae (Compositae: Lactuceae) and its taxonomic significance. Grana 21: 149–160.

Blackmore, S. 1984. The Northwest European Pollen Flora, 32: Compositae-Lactuceae. Rev. Palaeobot. Palynol. 42: 45–85.

Blackmore, S. 2000. The palynological compass: The contribution of palynology to systematics. Pp. 161–177 *in* B. Nordenstam, G. El-Ghazaly & M. Kassas (editors), Plant Systematics for the 21st Century. Portland Press, London.

Blackmore, S. 2007. Pollen and spores: Microscopic keys to understanding the earth's biodiversity. Pl. Syst. Evol. 263: 3–12.

Blackmore, S. & M. J. Cannon. 1983. Palynology and systematics of Morinaceae. Rev. Palaeobot. Palynol. 40: 207–226.

Blackmore, S. & P. R. Crane. 1988. Systematic implications of pollen and spore ontogeny. Pp. 83–115 *in* C. J. Humphries (editor), Ontogeny and Systematics. Columbia University Press, New York.

Blackmore, S. & S. H. Barnes. 1995. Garside's rule and the microspore tetrads of *Grevillea rosmarinifolia* A. Cunn. and *Dryandra polycephala* Benth. (Proteaceae). Rev. Palaeobot. Palynol. 85: 111–121.

Blackmore, S. & P. R. Crane. 1998. The evolution of apertures in the spores and pollen grains of embryophytes. Pp. 159–182 *in* S. J. Owens & P. J. Rudall (editors), Reproductive Biology. Royal Botanic Gardens, Kew, London.

Blackmore, S., A. H. Wortley, J. J. Skvarla & J. R. Rowley. 2007. Pollen wall development in flowering plants. New Phytol. 174: 482–498.

Blackmore, S., A. H. Wortley, J. J. Skvarla & H. Robinson. 2009. Evolution of pollen in the Compositae. Pp. 101–130 *in* V. A. Funk, A. Susanna, T. F. Stuessy & R. J. Bayer (editors), Systematics, Evolution and Biogeography of Compositae. IAPT, Vienna.

Blarer, A., D. L. Nickrent, H. Bänziger, P. K. Endress & Y.-L. Qiu. 2000. Phylogenetic relationships among genera of the parasitic family Rafflesiaceae s.lat. based on nuclear ITS and SSU rDNA, mitochondrial LSU and SSU rDNA, *atp1*, and *matR* sequences. Amer. J. Bot. 87: 503.

Blarer, A., D. L. Nickrent & P. K. Endress. 2004. Comparative floral structure and systematics in Apodanthaceae (Rafflesiales). Pl. Syst. Evol. 245: 119–142.

Borsch, T., C. Löhne & J. Wiersema. 2008. Phylogeny and evolutionary patterns in Nymphaeales: Integrating genes, genomes and morphology. Taxon 57: 1052–1081.

Bove, C. P. 1993. Pollen morphology of the Bignoniaceae from a south Brazilian Atlantic forest. Grana 32: 330–337.

Bremer, B., K. Bremer, N. Heidari, P. Erixon, R. G. Olmstead, A. A. Anderberg, M. Kallersjö & E. Barkhordarian. 2002. Phylogenetics of asterids based on 3 coding and 3 non-coding chloroplast DNA markers and the utility of non-coding DNA at higher taxonomic levels. Molec. Phylogen. Evol. 24: 274–301.

Bronckers, F. & F. Horvat. 1963. Note sur la morphologie du pollen de *Gossypium hirsutum* L. Pollen & Spores 5: 5–16.

Brown, R. 1811. On the Proteaceae of Jussieu. Trans. Linn. Soc. London 10: 15–226.

Brückner, P. 1993. Pollen morphology and taxonomy of Eurasiatic species of the genus *Buxus* (Buxaceae). Grana 32: 65–78.

Brusatte, S. L. 2010. Representing supraspecific taxa in higher-level phylogenetic analyses: Guidelines for palaeontologists. Palaeontology 53: 1–9.

Burns-Balogh, P. 1983. A theory on the evolution of the exine in Orchidaceae. Amer. J. Bot. 70: 1304–1312.

Cerceau-Larrival, M. T. & F. Roland-Heydacker. 1976. The evolutionary significance of the ultrastructure of the exine in Umbelliferous pollen grains. Pp. 481–498 *in* I. K. Ferguson & J. Muller (editors), The Evolutionary Significance of the Exine. Royal Botanic Gardens, Kew, London.

Chase, M. W., P. J. Rudall, M. F. Fay & K. L. Stobart. 2000. Xeronemataceae—A new family of Asparagoid lilies from New Caledonia and New Zealand. Kew Bull. 55: 865–870.

Chaturvedi, M. & K. Datta. 2001. Pollen morphology in *Saccharum* L. (Poaceae)—Wild and cultivated sugar cane species. Feddes Repert. 112: 387–390.

Chaturvedi, M., D. Yunus & K. Datta. 1994. Pollen morphology of *Sorghum* Moench—Sections *Eu-sorghum* and *Para-sorghum*. Grana 33: 117–123.

Chaturvedi, M., K. Datta & P. K. K. Nair. 1998. Pollen morphology of *Oryza* (Poaceae). Grana 37: 79–86.

Chinnappa, C. C. & B. G. Warner. 1981. Pollen morphology in the genus *Coffea* (Rubiaceae) and its taxonomic significance. Bot. J. Linn. Soc. 83: 221–236.

Chinnappa, C. C. & B. G. Warner. 1982. Pollen morphology in the genus *Coffea* (Rubiaceae). II. Pollen polymorphism. Grana 21: 29–37.

Christensen, J. E., H. T. Horner & N. R. Lersten. 1972. Pollen wall and tapetal orbicular wall development in *Sorghum bicolor* (Gramineae). Amer. J. Bot. 59: 43–58.

Chung, M. G. & S. B. Jones. 1989. Pollen morphology of *Hosta* Tratt. (Funkiaceae) and related genera. Bull. Torrey Bot. Club 116: 31–44.

Clarke, G. C. S., W. Punt & P. P. Hoen. 1991. The Northwest European Pollen Flora, 51: Ranunculaceae. Rev. Palaeobot. Palynol. 69: 117–271.

Claxton, F., H. I. Banks, B. B. Klitgaard & P. R. Crane. 2005. Pollen morphology of families Quillajaceae and Surianaceae (Fabales). Rev. Palaeobot. Palynol. 133: 221–233.

Cresti, M., F. Ciampolini, D. L. M. Mulcahy & G. Mulcahy. 1985. Ultrastructure of *Nicotiana alata* pollen, its germination and early tube formation. Amer. J. Bot. 72: 719–727.

Crompton, C. W. & W. F. Grant. 1983. Pollen morphology in Loteae (Leguminosae) with particular reference to the genus *Lotus* L. Grana 32: 129–153.

Cronquist, A. 1981. An Integrated System of Classification of Flowering Plants. Columbia University Press, New York.

Cronquist, A. 1988. The Evolution and Classification of Flowering Plants. New York Botanical Garden Press, Bronx, New York.

Cunningham, C. W., K. E. Omland & T. H. Oakley. 1998. Reconstructing ancestral character states: A critical reappraisal. Trends Ecol. Evol. 13: 361–366.

Dahlgren, G. 1988. An updated angiosperm classification. Bot. J. Linn. Soc. 100: 197–203.

Dahlgren, R. M. T. & H. T. Clifford. 1982. The Monocotyledons: A Comparative Study. Academic Press, London.

Danforth, B. R., S. Sipes, J. Fang & S. G. Brady. 2006. The history of early bee diversification based on five genes plus morphology. Proc. Natl. Acad. Sci. U.S.A. 103: 15,118–15,123.

de Queiroz, K. 1996. Including the characters of interest during tree reconstruction and the problems of circularity and bias in studies of character evolution. Amer. Naturalist 148: 700–708.

Dehgan, B. & N. B. Dehgan. 1988. Comparative pollen morphology and taxonomic affinities in Cycadales. Amer. J. Bot. 75: 1501–1516.

Denk, T. & M. V. Tekleva. 2006. Comparative pollen morphology and ultrastructure of *Platanus*: Implications for phylogeny and evaluation of the fossil record. Grana 45: 195–221.

Diethart, B., S. Sam & M. Weber. 2007. Walls of allergenic pollen: Special reference to the endexine. Grana 46: 164–175.

Dönmez, E. O. & S. Işik. 2008. Pollen morphology of Turkish Amaryllidaceae, Ixioliriaceae and Iridaceae. Grana 47: 15–38.

Donoghue, M. J. 1983. A preliminary analysis of phylogenetic relationships in *Viburnum* (Caprifoliaceae s.l.). Syst. Bot. 8: 45–58.

Donoghue, M. J. & J. A. Doyle. 1989. Phylogenetic analysis of angiosperms and the relationships of Hamamelidae. Pp. 17–45 *in* P. R. Crane & S. Blackmore (editors), Evolution, Systematics, and Fossil History of the Hamamelidae. Clarendon Press, Oxford.

Doyle, J. A. 2000. Paleobotany, relationships, and geographic history of Winteraceae. Ann. Missouri Bot. Gard. 87: 303–316.

Doyle, J. A. 2005. Early evolution of angiosperm pollen as inferred from molecular and morphological phylogenetic analyses. Grana 44: 227–251.

Doyle, J. A. 2009. Evolutionary significance of granular exine structure in the light of phylogenetic analyses. Rev. Palaeobot. Palynol. 156: 198–210.

Doyle, J. A. & C. L. Hotton. 1991. Diversification of early angiosperm pollen in a cladistic context. Pp. 69–195 *in* S. Blackmore & S. H. Barnes (editors), Pollen and Spores: Patterns of Diversification. Clarendon Press, Oxford.

Doyle, J. A., C. L. Hotton & J. V. Ward. 1990. Early Cretaceous tetrads, zonasulculate pollen, and Winteraceae II. Cladistic analysis and implications. Amer. J. Bot. 77: 1558–1568.

Drinnan, A. N., P. R. Crane & S. B. Hoot. 1994. Patterns of floral evolution in the early diversification of non-magnoliid dicotyledons (eudicots). Pl. Syst. Evol. 8: 93–122.

Dunbar, A. 1984. Pollen morphology in Campanulaceae IV. Nordic J. Bot. 4: 1–19.

Eklund, H., J. A. Doyle & P. S. Herendeen. 2004. Morphological phylogenetic analysis of living and fossil Chloranthaceae. Int. J. Pl. Sci. 165: 107–151.

Ekman, S., H. L. Andersen & M. Wedin. 2008. The limitations of ancestral state reconstruction and the evolution of the ascus in the Lecanorales (lichenized Ascomycota). Syst. Biol. 57: 141–156.

El-Ghazaly, G. & R. Chaudhary. 1993. Morphology and taxonomic application of orbicules (Ubisch bodies) in the genus *Euphorbia*. Grana 2: 26–32.

El-Ghazaly, G. A. & W. A. Jensen. 1986. Studies of the development of wheat (*Triticum aestivum*) pollen: 1. Formation of the pollen wall and Ubisch bodies. Grana 25: 1–29.

El-Ghazaly, G. A. & W. A. Jensen. 1990. Development of wheat (*Triticum aestivum*) pollen wall before and after effect of a gametocide. Canad. J. Bot. 68: 2509–2516.

Eliseu, S. A. & A. M. Dinis. 2008. Ultrastructure and cytochemistry of *Eucalyptus globulus* (Myrtaceae) pollen grain. Grana 47: 39–51.

Endress, P. K. & M. L. Matthews. 2006. First steps towards a floral structural characterization of the major rosid subclades. Pl. Syst. Evol. 260: 223–251.

Erdtman, G. 1943. An Introduction to Pollen Analysis. Ronald Press, New York.

Erdtman, G. 1944. Pollen morphology and plant taxonomy. II. Notes on some monocotyledonous pollen types. Svensk Bot. Tidskr. 38: 163–168.

Erdtman, G. 1945. Pollen morphology and plant taxonomy. IV. Labiatae, Verbenaceae, and Avicenniaceae. Svensk Bot. Tidskr. 39: 279–285.

Erdtman, G. 1946. Pollen morphology and plant taxonomy. VII. Notes on various families. Svensk Bot. Tidskr. 40: 77–84.

Erdtman, G. 1948. Pollen morphology and plant taxonomy. VIII. Didiereaceae. Bull. Mus. Natl. Hist. Nat., Sér. 3, Bot. 20: 387–394.

Erdtman, G. 1952. Pollen Morphology and Plant Taxonomy: Angiosperms (An Introduction to Palynology. I). Almqvist & Wiksell, Stockholm.

Erdtman, G. 1957. Pollen and Spore Morphology/Plant Taxonomy: Gymnospermae, Pteridophyta, Bryophyta. Almqvist & Wiksell, Stockholm.

Erdtman, G. 1960. The acetolysis method. A revised description. Svensk Bot. Tidskr. 54: 561–564.

Erdtman, G. 1971. On the affinities of *Aquilapollenites* Rouse. Grana 11: 15–17.

Erdtman, G., P. Leins, R. Melville & C. R. Metcalfe. 1969. On the relationships of *Emblingia*. Bot. J. Linn. Soc. 62: 169–186.

Esteban, R., J. M. Olano, J. Castresana, B. Fernández-Marín, A. Hernández, J. M. Beceriil & J. I. García-Plazaola. 2009. Distribution and evolutionary trends of photoprotective isoprenoids (xanthophylls and tocopherols) within the plant kingdom. Physiol. Pl. (Copenhagen) 135: 379–389.

Fægri, K. 1956. Recent trends in palynology. Bot. Rev. (Lancaster) 22: 639–664.

Fægri, K. & J. Iversen. 1989. Textbook of Pollen Analysis. Munksgaard, Copenhagen.

Ferguson, I. K. & J. Muller (editors). 1976. The Evolutionary Significance of the Exine. Academic Press Inc., London.

Ferguson, I. K. & J. J. Skvarla. 1981. The pollen morphology of the subfamily Papilionoideae (Leguminosae). Pp. 859–896 *in* R. M. Polhill & P. H. Raven (editors), Advances in Legume Systematics. Royal Botanic Gardens, Kew, London.

Ferguson, I. K. & M. M. Harley. 1993. The significance of new and recent work on pollen morphology in the Palmae. Kew Bull. 48: 205–225.

Filipowicz, N. & S. S. Renner. 2010. The worldwide holoparasitic Apodanthaceae confidently placed in the Cucurbitales by nuclear and mitochondrial genes. BMC Evol. Biol. 10: 1–8.

Freudenstein, J. V. & F. N. Rasmussen. 1999. What does morphology tell us about orchid relationships?—A cladistic analysis. Amer. J. Bot. 86: 225–248.

Funk, V. A. & T. F. Stuessy. 1978. Cladistics for the practicing plant taxonomist. Syst. Bot. 3: 159–178.

Funk, V. A., R. J. Bayer, S. C. Keeley, R. Chan, L. Watson, B. Gemeinholzer, E. E. Schilling, J. L. Panero, B. G. Baldwin, N. Garcia-Jacas, A. Susanna & R. K. Jansen. 2005. Everywhere but Antarctica: Using a supertree to understand the diversity and distribution of the Compositae. Biol. Skr. 55: 343–374.

Furness, C. A. & P. J. Rudall. 2003. Apertures with lids: Distribution and significance of operculate pollen in monocotyledons. Int. J. Pl. Sci. 164: 835–854.

Furness, C. A., S. Magallón & P. J. Rudall. 2007. Evolution of endoapertures in early-divergent eudicots, with particular reference to pollen morphology in Sabiaceae. Pl. Syst. Evol. 263: 77–92.

Godwin, H. 1968. The origin of the exine. New Phytol. 67: 667–676.

Godwin, H., P. Echlin & B. Chapman. 1967. The development of the pollen grain wall in *Ipomoea purpurea* (L.) Roth. Rev. Palaeobot. Palynol. 3: 181–195.

Goldberg, A. & H. A. Alden. 2005. Taxonomy of *Haptanthus* Goldberg & C. Nelson. Syst. Bot. 30: 773–778.

Graham, S. W. & R. G. Olmstead. 2000. Utility of 17 chloroplast genes for inferring the phylogeny of the basal angiosperms. Amer. J. Bot. 87: 1712–1730.

Grant, M., S. Blackmore & C. Morton. 2000. Pollen morphology of the subfamily Aurantioideae (Rutaceae). Grana 39: 8–20.

Grayum, M. H. 1992. Comparative external pollen ultrastructure of the Araceae and putatatively related taxa. Monogr. Syst. Bot. Missouri Bot. Gard. 43: 1–167.

Grew, N. 1682. The Anatomy of Plants. Rawline, London.

Gustafsson, M. H. G., E. Grafstrom & S. Nilsson. 1997. Pollen morphology of the Goodeniaceae and comparisons with related families. Grana 36: 185–207.

Hansen, A. K., L. E. Gilbert, B. B. Simpson, S. R. Downie, A. C. Cervi & R. K. Jansen. 2006. Phylogenetic relationships and chromosome number evolution in *Passiflora*. Syst. Bot. 31: 138–150.

Hansen, H. V. 1997. Studies in the Goodeniaceae and the Brunoniaceae with a discussion of their relationship to Asteraceae and Calyceraceae. Nordic J. Bot. 17: 495–510.

Hanson-Smith, V., B. Kolaczkowski & J. W. Thomas. 2010. Robustness of ancestral sequence reconstruction to phylogenetic uncertainty. Molec. Biol. Evol. 27: 1988–1999.

Harley, M. M. & J. Dransfield. 2003. Triporate pollen in the Arecaceae. Grana 42: 3–19.

Hawkins, B. A., M. A. Rodríguez & S. G. Weller. 2011. Global angiosperm family richness revisited: Linking ecology and evolution to climate. J. Biogeogr. 38: 1253–1266.

Horner, H. T. & C. B. Pearson. 1978. Pollen wall and aperture development in *Helianthus annuus* (Compositae: Heliantheae). Amer. J. Bot. 65: 293–309.

Huelsenbeck, J. P. & J. P. Bollback. 2001. Empirical and hierarchical Bayesian estimation of ancestral states. Syst. Biol. 50: 351–366.

Huelsenbeck, J. P., B. Rannala & J. P. Masly. 2000. Accommodating phylogenetic uncertainty in evolutionary studies. Science 288: 2349–2350.

Huysmans, S., G. El-Ghazaly & E. F. Smets. 1998. Orbicules in angiosperms: Morphology, function, distribution, and relation with tapetum tyoes. Bot. Rev. (Lancaster) 64: 240–272.

Inceoglu, O., N. M. Pinar & E. Oybak Donmez. 2000. Pollen morphology of wild *Vitis sylvestris* Gmelin (Vitaceae). Turkish J. Bot. 24: 147–150.

Jabbour, F., S. Nadot & C. Damerval. 2009. Evolution of floral symmetry: A state of the art. Compt. Rend. Biol. 332: 219–231.

Jansen, R. K., R. S. Wallace, K.-J. Kim & K. L. Chambers. 1991. Systematic implications of chloroplast DNA variation in the subtribe Microseridinae (Asteraceae: Lactuceae). Amer. J. Bot. 78: 1015–1027.

Jansen, R. K., Z.-Q. Cai, L. A. Raubeson, H. Daniel, C. W. de Pamphilis, J. Leebens-Mack, K. F. Müller, M. Guisinger-Bellian, R. C. Haberle, A. K. Hansen, T. W. Chumley, S.-B. Lee, R. Peery, J. R. McNeal, J. V. Kuehl & J. L. Boore. 2007. Analysis of 81 genes from 64 plastid genomes resolved relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc. Natl. Acad. Sci. U.S.A. 104: 19,369–19,374.

Jian, S., P. S. Soltis, M. A. Gitzendanner, M. J. Moore, R. Li, T. A. Hendry, Y.-L. Qiu, A. Dhingra, C. D. Bell & D. E. Soltis. 2008. Resolving an ancient, rapid radiation in Saxifragales. Syst. Biol. 57: 38–57.

Jones, M. R. & G. C. S. Clarke. 1981. The Northwest European Pollen Flora, 25: Nymphaeaceae. Rev. Palaeobot. Palynol. 33: NEPF 57–68.

Khan, R. 2005. Palynological studies of some genera from Brassicaceae. Int. J. Biol. Biotechnol. 2: 557–563.

Köhler, E. & E. Lange. 1979. A contribution to distinguishing cereal from wild grass pollen grains by LM and SEM. Grana 18: 133–140.

Köhler, E. & P. Brückner. 1982. Die pollenmorphologie der afrikanischen *Buxus*- und *Notobuxus*-arten (Buxaceae) und ihre systematische bedeutung. Grana 21: 71–82.

Koti, S., K. R. Reddy, V. G. Kakani, D. Zhao & V. R. Reddy. 2004. Soybean (*Glycine max*) pollen germination characteristics, flower and pollen morphology in response to enhanced ultraviolet-B radiation. Ann. Bot. (Oxford) 94: 855–864.

Kress, W. J. & D. E. Stone. 1983. Morphology and phylogenetic significance of exine-less pollen of *Heliconia* (Heliconiaceae). Syst. Bot. 8: 149–167.

Ladd, P. G. 1977. Pollen morphology of some members of the Restionaceae and related families with notes on the fossil record. Grana 16: 1–14.

Lahham, J. N. & D. Al-Eisawi. 1987. Pollen morphology of Jordanian Cruciferae. Mitt. Bot. Staatssamml. München 23: 355–375.

Lan, S.-Y., Z.-X. Xu, T.-D. Fu & W. K. Heneen. 1995. Pollen wall structure using a new stripping-sputtering device for scanning electron microscopy. Grana 34: 325–331.

Larson, D. A. 1966. On the significance of the detailed structure of *Passiflora caerulea* exines. Bot. Gaz. (London) 127: 40–48.

Le Thomas, A. & B. Lugardon. 1976. The granular structure of the columellar structure in pollen grains of Annonaceae. Adansonia 15: 543–572.

Lee, S. & E. J. Park. 1982. A cladistic analysis of the Korean Oleaceae. Korean J. Bot. 25: 57–64.

Lei, L.-G. & H.-X. Liang. 1998. Pollen morphology and its taxonomic significance of Piperaceae. Acta Bot. Yunnan. 20: 429–433.

Lesins, K. & I. Lesins. 1963. Pollen morphology and species relationship in *Medicago* L. Canad. J. Genet. Cytol. 5: 270–280.

Levin, R. A., W. L. Wagner, P. C. Hoch, M. Nepokroeff, J. C. Pires, E. A. Zimmer & K. J. Sytsma. 2003. Family-level relationships in the Onagraceae based on chloroplast *rbc*L and *ndh*F data. Amer. J. Bot. 90: 107–115.

Linder, H. P. 1984. A phylogenetic classification of the genera of the African Restionaceae. Bothalia 15: 11–76.

Liu, H. & C. S. Yang. 1989. Pollen morphology of Illiciaceae and its significance in systematics. Chin. J. Bot. 1: 104–115.

Lobreau-Callen, D. & M. S. Cervera. 1997. Pollen exine ultrastructure of the apetalous Crotonoideae. Rev. Palaeobot. Palynol. 98: 257–291.

Lundberg, J. & K. Bremer. 2003. A phylogenetic study of the order Asterales using one morphological and three molecular data sets. Int. J. Pl. Sci. 164: 553–578.

Lu, L., A. H. Wortley, D.-Z. Li, H. Wang & S. Blackmore. 2015. Evolution of Angiosperm Pollen: 2. The Basal Angiosperms. Ann. Missouri Bot. Gard. 100(3–4): 227–269.

Mabberley, D. J. 2008. Mabberley's Plant Book. A Portable Dictionary of Plants, Their Classification and Uses. Cambridge University Press, Cambridge.

Maddison, D. R. & W. P. Maddison. 2000. MacClade 4: Analysis of Phylogeny and Character Evolution. Sinauer Associates, Sunderland, Massachusetts.

Maddison, W. P. & D. R. Maddison. 2009. Mesquite: A Modular System for Evolutionary Analysis, Version 2.6.

Maguire, B., J. J. Wurdack & Y.-C. Huang. 1974. Pollen grains of some American Olacaceae. Grana 14: 26–38.

Martins, L. H. P., I. P. A. Miranda & C. D. Nunes. 2003. Pollen morphology of Amazonian populations of *Elaeis oleifera*. Acta Amazonica 33: 159–166.

Mathew, P. J. & P. M. Mathew. 2001. Pollen morphology of some members of Pipereaceae and its bearing on the systematics and phylogeny of the family. Rheedea 11: 65–78.

Matthews, M. L. & P. K. Endress. 2004. Comparative floral structure and systematics in Cucurbitales (Corynocarpaceae, Coriariaceae, Tetramelaceae, Datiscaceae, Begoniaceae, Cucurbitaceae, Anisophylleaceae). Bot. J. Linn. Soc. 145: 129–185.

Mbagwu, F. N., V. U. Okafor & A. E. Okolo. 2008. Comparative palynological characters on four species of the family Poaceae (Gramineae). Int. Sci. Res. J. 1: 127–129.

Médus, J., R. Gajardo & P. Woltz. 1988. Exine ultrastructure of *Dacrydium fonkii*, *Saxegothaea conspicua* and *Stachycarpus andina* (Podocarpaceae) from southern South America. Grana 28: 19–23.

Merckx, V., P. Schols, K. Geuten, S. Huysmans & E. F. Smets. 2008. Phylogenetic relationships in Nartheciaceae (Dioscoreales), with focus on pollen and orbicule morphology. Belg. J. Bot. 141: 64–77.

Millay, M. A. & T. N. Taylor. 1976. Evolutionary trends in fossil gymnosperm pollen. Rev. Palaeobot. Palynol. 21: 65–91.

Mooers, A. O. & D. Schluter. 1999. Reconstructing ancestor states with maximum likelihood: Support for one- and two-rate models. Syst. Biol. 48: 623–633.

Moore, M. J., C. D. Bell, P. S. Soltis & D. E. Soltis. 2007. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. Proc. Natl. Acad. Sci. U.S.A. 104: 19,363–19,368.

Moore, M. J., P. S. Soltis, C. D. Bell, J. G. Burleigh & D. E. Soltis. 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. Proc. Natl. Acad. Sci. U.S.A. 107: 4623–4628.

Muller, J. 1970. Palynological evidence on early differentiation of angiosperms. Biol. Rev. Cambridge Philos. Soc. 45: 417–450.

Muller, J. 1981. Fossil pollen records of extant angiosperms. Bot. Rev. (Lancaster) 47: 1–142.

Muller, J. 1984. Sigificance of fossil pollen for angiosperm history. Ann. Missouri Bot. Gard. 71: 419–443.

Nickrent, D. L., A. Blarer, Y.-L. Qiu, V. Vidal-Russel & F. E. Anderson. 2004. Phylogenetic inference in Rafflesiales: The influence of rate heterogeneity and horizontal gene transfer. BMC Evol. Biol. 4: 40.

Nickrent, D. L., J. P. Der & F. E. Anderson. 2005. Discovery of the photosynthetic relatives of the "Maltese mushroom" *Cynomorium*. BMC Evol. Biol. 5: 38.

Nielsen, R. 2002. Mapping mutations on phylogenies. Syst. Biol. 51: 729–739.

Nilsson, S., J. Praglowski & L. Nilsson. 1977. Atlas of Airborne Pollen Grains and Spores in Northern Europe. Natur och Kultur, Stockholm.

Nilsson, S., J. Coetzee & E. Grafstrom. 1996. On the origin of the Sarcolaenaceae with reference to pollen morphological evidence. Grana 35: 321–334.

Nowicke, J. W. 1994. A palynological study of Crotonoideae (Euphorbiaceae). Ann. Missouri Bot. Gard. 81: 245–269.

Nowicke, J. W. & J. J. Skvarla. 1979. Pollen morphology: The potential influence in higher order systematics. Ann. Missouri Bot. Gard. 66: 633–700.

Nowicke, J. W. & J. J. Skvarla. 1981. Pollen morphology and phylogenetic relationships of the Berberidaceae. Smithsonian Contr. Bot. 50: 1–83.

Nowicke, J. W. & J. J. Skvarla. 1984. Pollen morphology and the relationships of *Simmondsia chinensis* to the order Euphorbiales. Amer. J. Bot. 71: 210–215.

Nyananyo, B. L. 1992. Pollen morphology in the Portulacaceae (Centrospermae). Folia Geobot. Phytotax. 27: 387–400.

Olvera, H. F., S. Fuentes-Soriano & E. M. Hernández. 2006. Pollen morphology and systematics of Atripliceae (Chenopodiaceae). Grana 45: 175–194.

Omland, K. E. 1999. The assumptions and challenges of ancestral state reconstructions. Syst. Biol. 48: 604–611.

Pacini, E., G. G. Franchi & M. Ripaccioli. 1999. Ripe pollen structure and histochemistry of some gymnosperms. Pl. Syst. Evol. 217: 81–99.

Padgett, D. J. 2007. A monograph of *Nuphar* (Nymphaeaceae). Rhodora 109: 1–95.

Pagel, M. 1997. Inferring evolutionary processes from phylogenies. Zool. Scripta 26: 331–348.

Pagel, M. 1999a. Inferring the historical patterns of biological evolution. Nature 401: 877–884.

Pagel, M. 1999b. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. Syst. Biol. 48: 612–622.

Pagel, M. & A. Meade. 2006. BayesTraits. University of Reading, Reading, United Kingdom.

Pagel, M., A. Meade & D. Barker. 2004. Bayesian estimation of ancestral character states on phylogenies. Syst. Biol. 53: 673–684.

Patel, V. C., J. J. Skvarla & P. H. Raven. 1984. Pollen characters in relation to the delimitation of Myrtales. Ann. Missouri Bot. Gard. 71: 858–969.

Patterson, C. 1982. Morphological characters and homology. Pp. 21–74 *in* K. A. Joysey & A. E. Friday (editors), Problems in Phylogenetic Reconstruction. Academic Press, London.

Pavón, J. 1798–1802. Flora Peruviana, et Chilensis: Sive, Descriptiones, et Icones Plantarum Peruvianarum, et Chilensium, Secundum Systema Linnaeanum Digestae, cum Characteribus Plurium Generum Evulgatorum Reformatis. Typis Gabrielis de Sancha, Madrid.

Pehlivan, S. 1987. A comparative study of the fine structures of the pollen walls and annuli in some Turkish Betulaceae, Moraceae, Cannabaceae, Haloragaceae. Commun. Fac. Sci. Univ. Ankara, Ser. C2, Botanique 5: 1–18.

Perveen, A. & M. Qaiser. 1998a. Pollen flora of Pakistan - X. Leguminosae (subfamily: Caesalpinioideae). Turkish J. Bot. 22: 145–150.

Perveen, A. & M. Qaiser. 1998b. Pollen flora of Pakistan XI. Leguminosae (subfamily: Mimosoideae). Turkish J. Bot. 22: 151–156.

Perveen, A. & M. Qaiser. 1998c. Pollen flora of Pakistan - VIII. Leguminosae (subfamily: Papilionoideae). Turkish J. Bot. 22: 73–91.

Perveen, A. & M. Qaiser. 1999. Pollen flora of Pakistan - XIII. Campanulaceae. Turkish J. Bot. 23: 45–51.

Perveen, A. & M. Qaiser. 2008a. Pollen flora of Pakistan - LVII. Vitaceae. Pakistan J. Bot. 40: 501–506.

Perveen, A. & M. Qaiser. 2008b. Pollen flora of Pakistan - LVI. Cucurbitaceae. Pakistan J. Bot. 40: 9–16.

Perveen, A., M. Qaiser & R. Khan. 2004. Pollen flora of Pakistan - XLII. Brassicaceae. Pakistan J. Bot. 36: 683–700.

Pike, K. M. 1955. Pollen morphology of Myrtaceae from the south-west Pacific area. Austral. J. Bot. 4: 13–54.

Polevova, S. V. 2006. Review of the sporoderm ultrastructure of members of the Asterales. Paleontol. Zurn. 40: S656–S663.

Praglowski, J. 1979. Angiospermae: Winteraceae Lindl. World Pollen Spore Fl. 8: 1–25.

Presting, D. 1965. On the morphology of the pollen grains of the Passifloraceae. Pollen & Spores 7: 193–247.

Punt, W. 1962. Pollen morphology of the Euphorbiaceae with special reference to taxonomy. Wentia 7: 1–116.

Punt, W. 1975. The Northwest European Pollen Flora, 5: Sparganiaceae and Typhaceae. Rev. Palaeobot. Palynol. 19: 75–88.

Punt, W. 1984. The Northwest European Pollen Flora, 37: Umbelliferae. Rev. Palaeobot. Palynol. 42: NEPF 155–364.

Punt, W. & M. Monnabrands. 1977. The Northwest European Pollen Flora, 8: Solanaceae. Rev. Palaeobot. Palynol. 23: 1–30.

Punt, W. & M. Malotaux. 1984. The Northwest European Pollen Flora, 31: Cannabaceae, Moraceae and Urticaceae. Rev. Palaeobot. Palynol. 42: 23–44.

Punt, W. & A. Marks. 1991. The Northwest European Pollen Flora, 50: Buxaceae. Rev. Palaeobot. Palynol. 69: 113–115.

Punt, W., J. A. A. Bos & P. P. Hoen. 1991. The Northwest European Pollen Flora, 45: Oleaceae. Rev. Palaeobot. Palynol. 69: 23–47.

Punt, W., S. Blackmore, S. Nilsson & A. Le Thomas (editors). 1994. Glossary of Pollen and Spore Terminology. LPP Foundation, Utrecht.

Punt, W., A. Marks & P. P. Hoen. 2003a. The Northwest European Pollen Flora, 64: Vitaceae. Rev. Palaeobot. Palynol. 123: 67–70.

Punt, W., J. Rovers & P. P. Hoen. 2003b. The Northwest European Pollen Flora, 67: Onagraceae. Rev. Palaeobot. Palynol. 123: 107–161.

Punt, W., P. P. Hoen, S. Blackmore, L. Nilsson & A. Le Thomas. 2007. Glossary of pollen and spore terminology. Rev. Palaeobot. Palynol. 143: 1–81.

Qiu, Y.-L., J. H. Lee, F. Bernasconi-Quadroni, D. E. Soltis, P. S. Soltis, M. J. Zanis, E. A. Zimmer, Z.-D. Chen, V.

Savolainen & M. W. Chase. 1999. The earliest angiosperms: Evidence from mitochondrial, plastid and nuclear genomes. Nature 402: 404–407.
Remizowa, M. V., D. D. Sokoloff, T. D. Macfarlane, S. R. Yadav, C. J. Prychid & P. J. Rudall. 2008. Comparative pollen morphology in the early-divergent angiosperm family Hydatellaceae reveals variation at the infraspecific level. Grana 47: 81–100.
Ronquist, F. 2004. Bayesian inference of character evolution. Trends Ecol. Evol. 19: 475–481.
Roquet, C., L. Sáez, J. J. Aldasoro, A. Susanna, M. L. Alarcón & N. Garcia-Jacas. 2008. Natural delimitation, molecular phylogeny and floral evolution in *Campanula*. Syst. Bot. 33: 203–217.
Roubik, D. W. & J. E. Moreno. 1991. Pollen and spores of Barro Colorado Island. Monogr. Syst. Bot. Missouri Bot. Gard. 36: 1–270.
Rowley, J. R. 1967. Fibrils, microtubules and lamellae in pollen grain. Rev. Palaeobot. Palynol. 3: 213–226.
Rowley, J. R. & G. Erdtman. 1967. Sporoderm in *Populus* and *Salix*. Grana 7: 517–567.
Rudall, P. J., C. A. Furness, M. W. Chase & M. F. Fay. 1997. Microsporogenesis and pollen sulcus type in Asparagales (Lilianae). Canad. J. Bot. 75: 408–430.
Saad, S. I. 1964. Pollen morphology of some Egyptian Cucurbitaceae. Pollen & Spores 6: 113–124.
Sage, T. L., K. Hristova-Sarkovski, V. Koehl, J. Lyew, V. Pontieri, P. Bernhardt, P. Weston, S. Bagha & G. Chiu. 2009. Transmitting tissue architecture in basal-relictual angiosperms: Implications for transmitting tissue origins. Amer. J. Bot. 96: 183–206.
Sahashi, N. & J. Ueno. 1986. Pollen morphology of *Ginkgo biloba* and *Cycas revoluta*. Canad. J. Bot. 64: 3075–3078.
Sampson, F. B. 1993. Pollen morphology of the Amborellaceae and Hortoniaceae (Hortonioideae, Monimiaceae). Grana 32: 154–162.
Sampson, F. B. 2000. Pollen diversity in some modern magnoliids. Int. J. Pl. Sci. 161: S193–S210.
Santos, F. H. 2010. Pollen morphology of *Arabidopsis thaliana* (L.) Heynh. (Brassicaceae) *in* Botania Seminalis (blog). <http://botaniaseminalis.blogspot.co.uk>, accessed 10 November 2010.
Sauquet, H. & D. J. Cantrill. 2007. Pollen diversity and evolution in Proteoideae (Proteales: Proteaceae). Syst. Bot. 32: 271–316.
Savolainen, V., M. W. Chase, S. B. Hoot, C. M. Morton, D. E. Soltis, C. Bayer, M. F. Fay, A. Y. De Bruijn, S. Sullivan & Y. L. Qiu. 2000. Phylogenetics of flowering plants based on combined analysis of plastid *atpB* and *rbcL* gene sequences. Syst. Biol. 49: 306–362.
Schill, R. & W. Pfeiffer. 1977. Untersuchungen an Orchideenpollinien unter besonderer Beruecksichtigung ihrer Feinskulpturen. Pollen & Spores 19: 5–118.
Schols, P., C. A. Furness, P. Wilkin & E. F. Smets. 2001. Morphology of pollen and orbicules in some *Dioscorea* species and its systematic implications. Bot. J. Linn. Soc. 136: 295–311.
Schols, P., C. A. Furness, P. Wilkin, E. F. Smets, V. Cielen & S. Huysmans. 2003. Pollen morphology of *Dioscorea* (Dioscoreaceae) and its relation to systematics. Bot. J. Linn. Soc. 143: 375–390.
Schols, P., C. A. Furness, V. Merckx, P. Wilkin & E. F. Smets. 2005. Comparative pollen development in Dioscoreales. Int. J. Pl. Sci. 166: 909–924.
Scotland, R. W. 1992. Pollen morphology and taxonomic characters in Acanthaceae. Syst. Bot. 17: 337–340.
Sengupta, S. 1972. On the pollen morphology of Convolvulaceae with special reference to taxonomy. Rev. Palaeobot. Palynol. 13: 157–212.
Sharma, B. D. 1974. Contribution to the palynotaxonomy of genus *Solanum* Linn. J. Palynol. 10: 51–68.
Sharma, M. 1967. Pollen morphology of Indian monocotyledons. J. Palynol. Special volume: 1–96.
Shehata, A. A. 2008. Pollen morphology of Egyptian Geraniaceae: An assessment of taxonomic value. Int. J. Bot. 4: 67–76.
Simmons, M. P. 2004. Celastraceae. Pp. 29–64 *in* K. Kubitzki (editor), The Families and Genera of Vascular Plants. VI. Flowering Plants. Dicotyledons. Celastrales, Oxalidales, Rosales, Cornales, Ericales. Springer, Berlin.
Skvarla, J. J. & D. A. Larson. 1966. Fine structural studies of *Zea mays* pollen I: Cell membranes and exine ontogeny. Amer. J. Bot. 53: 1112–1125.
Skvarla, J. J., P. H. Raven & J. Praglowski. 1976. Ultrastructural survey of Onagraceae pollen. Pp. 447–480 *in* I. K. Ferguson & J. Muller (editors), The Evolutionary Significance of the Exine. Academic Press, London.
Skvarla, J. J., P. H. Raven, W. F. Chissoe & M. Sharp. 1978. An ultrastructural study of viscin threads in Onagraceae pollen. Pollen & Spores 20: 5–143.
Soltis, D. E., P. S. Soltis, D. R. Morgan, S. M. Swensen, B. C. Mullin, J. M. Dowd & P. G. Martin. 1995. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. Proc. Natl. Acad. Sci. U.S.A. 92: 2647–2651.
Soltis, D. E., P. S. Soltis & M. W. Chase. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. Nature 402: 402–404.
Soltis, D. E., P. S. Soltis, M. W. Chase, M. E. Mort, D. C. Albach, M. J. Zanis, V. Savolainen, W. H. Hahn, S. B. Hoot, M. F. Fay, M. Axtell, S. M. Swensen, L. M. Prince, W. J. Kress, K. C. Nixon & J. S. Farris. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL* and *atpB* sequences. Bot. J. Linn. Soc. 133: 381–461.
Soltis, D. E., M. A. Gitzendanner & P. S. Soltis. 2007. A 567-taxon data set for angiosperms: The challenges posed by Bayesian analyses of large data sets. Int. J. Pl. Sci. 168: 137–157.
Soltis, D. E., C. D. Bell, S. Kim & P. S. Soltis. 2008. Origin and early evolution of angiosperms. Ann. New York Acad. Sci. 1133: 3–25.
Soltis, D. E., S. A. Smith, N. Cellinese, K. J. Wurdack, D. C. Tank, S. F. Brockington, N. F. Refulio-Rodriguez, J. B. Walker, M. J. Moore, B. S. Carlsward, C. D. Bell, M. Latvis, S. Crawley, C. Black, D. Diouf, Z.-X. Xi, C. A. Rushworth, M. A. Gitzendanner, K. J. Sytsma, Y.-L. Qiu, K. W. Hilu, C. C. Davis, M. J. Sanderson, R. S. Beaman, R. G. Olmstead, W. S. Judd, M. J. Donoghue & P. S. Soltis. 2011. Angiosperm phylogeny: 17 genes, 640 taxa. Amer. J. Bot. 98: 704–730.
Sowunmi, M. A. 1972. Pollen morphology of Palmae and its bearing on taxonomy. Rev. Palaeobot. Palynol. 13: 1–80.
Stafford, P. J. & M. Gibby. 1992. Pollen morphology of the genus *Pelargonium* (Geraniaceae). Rev. Palaeobot. Palynol. 71: 79–109.
Steindl, F. 1945. Beitrag zur Pollen und Embryobil-dung bei *Cynomorium coccineum* L. Arch. Julius Klaus Stiftung 20: 342–355.
Stoffelen, P., E. Robbrecht & E. F. Smets. 1997. Pollen morphology of *Coffea* and *Psilanthus* (Rubiaceae-Coffeeae), mainly from Africa. Grana 36: 313–327.

Suzuki, T., K. Masaoka, M. Nishi, K. Nakamura & S. Ishiguro. 2008. Identification of *kaonashi* mutants showing abnormal pollen exine structure in *Arabidopsis thaliana*. Pl. Cell Physiol. 49: 1465–1477.

Takahashi, M. 1994. Exine development in *Illicium religiosum* Sieb. et Zucc. (Illiciaceae). Grana 33: 309–312.

Takhtajan, A. L. 1980. Outline of the classification of flowering plants (Magnoliophyta). Bot. Rev. (Lancaster) 46: 225–359.

Takhtajan, A. L., N. R. Meyer & V. N. Kosenko. 1985. Pollen morphology and classification in Rafflesiaceae s.l. Bot. Zhurn. (Moscow & Leningrad) 70: 153–162.

Tellería, M. C. & G. Daners. 2003. Pollen types in southern New World Convolvulaceae and their taxonomic significance. Pl. Syst. Evol. 243: 99–118.

Terekhin, E. S., M. S. Yakovlev & Z. I. Nikiticheva. 1975. Development of the microsporangium, pollen grains, ovule and embryosac in *Cynomorium songaricum* (Cynomoriaceae). Bot. Zhurn. (Moscow & Leningrad) 60: 153–162.

Thorne, R. F. 1992. An updated phylogenetic classification of flowering plants. Aliso 13: 365–389.

Tomb, A. S. 1999. Pollen morphology and relationships of *Setchellanthus caeruleus* (Setchellanthaceae). Taxon 48: 285–288.

Torices, R. 2010. Adding time-calibrated branch lengths to the Asteraceae supertree. J. Syst. Evol. 48: 271–278.

Trigo, M. 1992. Contribution to the study of pollen in ornamental species: Solanaceae, Convolvulaceae and Hydrophyllaceae. Acta Bot. Malac. 17: 209–222.

Vaissière, B. E. & S. B. Vinson. 1994. Pollen morphology and its effect on pollen collection by honey-bees, *Apis mellifera* L. (Hymenoptera, Apidae), with special reference to upland cotton, *Gossypium hirsutum* L. (Malvaceae). Grana 33: 128–138.

Van Campo, E. & B. Lugardon. 1973. Structure grenue infratectale de l'ectexine des pollens de quelques Gymnospermes et Angiospermes. Comp. Rend. Acad. Sci. Paris, Sér. D, 272: 2071–2074.

Van Campo, M. 1976. Patterns of pollen morphological variation within taxa. Pp. 125–137 *in* I. K. Ferguson & J. Mueller (editors), The Evolutionary Significance of the Exine. Academic Press, London.

Vanderpoorten, A. & B. Goffinet. 2006. Mapping uncertainty and phylogenetic uncertainty in ancestral character states reconstruction: An example in the moss genus *Brachytheciastrum*. Syst. Biol. 55: 957–971.

Vinckier, S., S. Huysmans & E. F. Smets. 2000. Morphology and ultrastructure of orbicules in the subfamily Ixoroideae (Rubiaceae). Rev. Palaeobot. Palynol. 108: 151–174.

von Mohl, H. 1835. Sur la structure et les formes des graines de pollen. Ann. Sci. Nat. (Paris), Ser. 2, 3: 148–180, 220–236, 304–346.

Walker, J. W. 1971. Unique type of angiosperm pollen from family Annonaceae. Science 172: 565.

Walker, J. W. 1974. Aperture evolution in pollen of primitive angiosperms. Amer. J. Bot. 61: 1112–1137.

Walker, J. W. 1975. The bases of angiosperm phylogeny: Introduction. Ann. Missouri Bot. Gard. 62: 515–516.

Walker, J. W. 1976. Evolutionary significance of the exine in the pollen of primitive angiosperms. Pp. 251–308 *in* D. K. Ferguson & J. Muller (editors), The Evolutionary Significance of the Exine. Academic Press, London.

Walker, J. W. & J. A. Doyle. 1975. The bases of angiosperm phylogeny: Palynology. Ann. Missouri Bot. Gard. 62: 664–723.

Walker, J. W. & A. G. Walker. 1984. Ultrastructure of Lower Cretaceous angiosperm pollen and the origin and early evolution of flowering plants. Ann. Missouri Bot. Gard. 71: 464–521.

Wang, F.-X., N.-F. Qian, Y.-L. Zang & H.-Q. Yang. 1995. Pollen Flora of China. Science Press, Beijing.

Wang, H.-C., M. J. Moore, P. S. Soltis, C. D. Bell, S. F. Brockington, R. Alexandre, C. C. Davis, M. Latvis, S. R. Manchester & D. E. Soltis. 2009a. Rosid radiation and the rapid rise of angiosperm-dominated forests. Proc. Natl. Acad. Sci. U.S.A. 106: 3853–3858.

Wang, H., H.-J. He, J.-Q. Chen & L. Lu. 2009b. Palynological data on Illiciaceae and Schisandraceae confirm phylogenetic relationships within these two basally-branching angiosperm families. Flora 205: 221–228.

Wanntorp, L., J. Praglowski & E. Grafstrom. 2004. New insights into the pollen morphology of the genus *Gunnera* (Gunneraceae). Grana 43: 15–21.

Wei, Z.-X. & Z.-Y. Wu. 1993. Pollen ultrastructure of *Liriodendron* and its systematic significance. Acta Bot. Yunnan. 15: 163–166.

Wen, J. & J. W. Nowicke. 1999. Pollen ultrastructure of Panax (the ginseng genus, Araliaceae), an eastern Asian and eastern North American disjunct genus. Amer. J. Bot. 86: 1624–1636.

Wiens, J. J., C. A. Kuczynski, W. E. Duellman & T. W. Reeder. 2007. Loss and re-evolution of complex life cycles in marsupial frogs: Does ancestral trait reconstruction mislead? Evolution 61: 1886–1899.

Wodehouse, R. P. 1926. Pollen grain morphology in the classification of the Anthemideae. Bull. Torrey Bot. Club 53: 479–485.

Wodehouse, R. P. 1928a. Pollen grains in the identification and classification of plants. II. *Barnadesia*. Bull. Torrey Bot. Club 55: 449–462.

Wodehouse, R. P. 1928b. Pollen grains in the identification and classification of plants. I. The Ambrosiaceae. Bull. Torrey Bot. Club 55: 181–198.

Wodehouse, R. P. 1928c. The phylogenetic value of pollen grain characters. Ann. Bot. (Oxford) 42: 891–934.

Wodehouse, R. P. 1929a. Pollen grains in the identification and classification of plants. III. The Nassauvinae. Bull. Torrey Bot. Club 56: 123–138.

Wodehouse, R. P. 1929b. Pollen grains in the identification and classification of plants. IV. The Mutisieae. Amer. J. Bot. 16: 297–313.

Wodehouse, R. P. 1930. Pollen grains in the identification and classification of plants. V. *Haplopappus* and other Astereae: The origin of their furrow configurations. Amer. J. Bot. 16: 297–313.

Wodehouse, R. P. 1935. Pollen Grains: Their Structure, Identification and Significance in Science and Medicine. McGraw-Hill, New York.

Wodehouse, R. P. 1936. Pollen grains in the identification and classification of plants VIII. The Alismataceae. Amer. J. Bot. 23: 535–539.

Worberg, A., M. H. Alford, D. Quandt & T. Borsch. 2009. Huerteales sister to Brassicales plus Malvales, and newly circumscribed to include *Dipentodon*, *Gerrardina*, *Huertea*, *Perrottetia*, and *Tapiscia*. Taxon 58: 468–478.

Wu, Z.-Y., P. H. Raven & D.-Y. Hong. 2010. Flora of China, Vol. 23 (Acoraceae through Cyperaceae). Science

Press, Beijing, and Missouri Botanical Garden Press, St. Louis.
Wurdack, K. J. & C. C. Davis. 2009. Malpighiales phylogenetics: Gaining ground on one of the most recalcitrant clades in the angiosperm tree of life. Amer. J. Bot. 96: 1551–1570.
Xiang, Q.-Y. & D. T. Thomas. 2008. Tracing character evolution and biogeographic history through time in Cornaceae—Does choice of methods matter? J. Syst. Evol. 46: 349–374.
Yang, Z.-H. 2006. Computational Molecular Evolution. Oxford University Press, Oxford.
Yao, Y.-F., Y.-Z. Xi, B.-Y. Geng & C.-S. Li. 2004. The exine ultrastructure of pollen grains in *Gnetum* (Gnetaceae) from China and its bearing on the relationship with the ANITA group. Bot. J. Linn. Soc. 146: 415–425.
Yeates, D. K. 1995. Groundplans and exemplars: Paths to the tree of life. Cladistics 11: 343–357.
Zavada, M. S. 1983. Comparative morphology of monocot pollen and evolutionary trends of apertures and wall structures. Bot. Rev. (Lancaster) 49: 331–379.
Zavada, M. S. 1990. A contribution to the study of pollen wall ultrastructure of orchid pollinia. Ann. Missouri Bot. Gard. 77: 785–801.
Zhang, Z.-M., K.-M. Cui & Z.-L. Li. 2000. Morphology and lateral germination of pollen in *Ginkgo biloba* and their implications in evolution. Acta Phytotax. Sin. 38: 141–147.
Zhang, Z.-Y., D.-Z. Yang, L. Joongku & L.-Q. Li. 2009. Supplemental study on the pollen morphology of the tribe Hyoscyameae (Solanaceae) and its systematic significance. Guihaia 29: 285–295.
Zhou, Q. & D. Fu. 2008. Reproductive morphology of *Nuphar* (Nymphaeaceae), a member of basal angiosperms. Pl. Syst. Evol. 272: 79–96.
Zhu, X.-Y., M. W. Chase, Y.-L. Qiu, H.-Z. Kong, D. L. Dilcher, J.-H. Li & Z.-D. Chen. 2007. Mitochondrial *matR* sequences help to resolve deep phylogenetic relationships in rosids. BMC Evol. Biol. 7: 1–15.

APPENDIX 1. Literature sources used for palynological data, with angiosperms listed alphabetically by taxonomic order.

Gymnosperm outgroups (Wodehouse, 1935; Erdtman, 1943, 1957; Walker, 1974; Millay & Taylor, 1976; Nilsson et al., 1977; Audran & Masure, 1978; Sahashi & Ueno, 1986; Dehgan & Dehgan, 1988; Médus et al., 1988; Pacini et al., 1999; Zhang et al., 2000).

Angiosperm orders. ACORALES (Erdtman, 1952; Grayum, 1992; Rudall et al., 1997; Furness & Rudall, 2003; Sage et al., 2009; Wu et al., 2010); AMBORELLALES (Erdtman, 1952; Walker, 1976; Dahlgren & Clifford, 1982; Sampson, 1993, 2000; Doyle, 2009); APIALES (Cerceau-Larrival & Roland-Heydacker, 1976; Punt, 1984; Huysmans et al., 1998; Wen & Nowicke, 1999); ARECALES (Sowunmi, 1972; Ferguson & Harley, 1993; Martins et al., 2003); ASPARAGALES (Erdtman, 1952; Schill & Pfeiffer, 1977; Burns-Balogh, 1983; Alvarez & Köhler, 1987; Chung & Jones, 1989; Zavada, 1990; Freudenstein & Rasmussen, 1999; Chase et al., 2000); ASTERALES (Erdtman, 1952; Horner & Pearson, 1978; Blackmore, 1984; Dunbar, 1984; Dehgan & Dehgan, 1988; Vaissière & Vinson, 1994; Gustafsson et al., 1997; Hansen, 1997; Perveen & Qaiser, 1999; Lundberg & Bremer, 2003; Polevova, 2006); AUSTROBAILEYALES (Erdtman, 1952; Liu & Yang, 1989; Takahashi, 1994; Sampson, 2000; Yao et al., 2004; Wang et al., 2009b); BRASSICALES (Erdtman et al., 1969; Nilsson et al., 1977; Lahham & Al-Eisawi, 1987; Tomb, 1999; Perveen et al., 2004; Khan, 2005; Diethart et al., 2007; Suzuki et al., 2008; Santos, 2010); BUXALES (Köhler & Brückner, 1982; Nowicke & Skvarla, 1984; Punt & Marks, 1991; Brückner, 1993; Goldberg & Alden, 2005); CANELLALES (Erdtman, 1952; Praglowski, 1979; Acosta Castellanos & Vilela, 1998; Doyle, 2000; Sampson, 2000); CARYOPHYLLALES (Wodehouse, 1935; Erdtman, 1952; Nyananyo, 1992; Behnke, 1997; Olvera et al., 2006); CHLORANTHALES (Walker & Walker, 1984; Eklund et al., 2004); CUCURBITALES (Erdtman, 1952; Saad, 1964; Matthews & Endress, 2004; Perveen & Qaiser, 2008a); DIOSCOREALES (Erdtman, 1952; Zavada, 1983; Schols et al., 2003, 2005); FABALES (Lesins & Lesins, 1963; Ferguson & Skvarla, 1981; Crompton & Grant, 1983; Perveen & Qaiser, 1998a, 1998b, 1998c; Koti et al., 2004; Claxton et al., 2005); GENTIANALES (Chinnappa & Warner, 1981, 1982; Stoffelen et al., 1997; Vinckier et al., 2000); GERANIALES (Larson, 1966; Stafford & Gibby, 1992; Shehata, 2008); LAMIALES (Erdtman, 1952; Lee & Park, 1982; Punt et al., 1991; Bove, 1993); LAURALES (Erdtman, 1952; Walker, 1976; Sampson, 2000); MAGNOLIALES (Wodehouse, 1935; Erdtman, 1952; Walker, 1974, 1976; Wei & Wu, 1993); MALPIGHIALES (Erdtman, 1952; Presting, 1965; Larson, 1966; Rowley & Erdtman, 1967; Nilsson et al., 1977; Nowicke, 1994; Lobreau-Callen & Cervera, 1997; Amela García et al., 2002; Hansen et al., 2006; Wurdack & Davis, 2009); MALVALES (Bronckers & Horvat, 1963; Beasley, 1975; Vaissière & Vinson, 1994; Lan et al., 1995; Nilsson et al., 1996); MYRTALES (Pike, 1955; Skvarla et al., 1976, 1978; Patel et al., 1984; Levin et al., 2003; Punt et al., 2003b; Eliseu & Dinis, 2008); NYMPHAEALES (Wodehouse, 1935; Erdtman, 1943, 1952; Rowley, 1967; Jones & Clarke, 1981; Sampson, 2000; Doyle, 2005; Padgett, 2007; Borsch et al., 2008; Remizowa et al., 2008; Zhou & Fu, 2008); PIPERALES (Erdtman, 1952; Lei & Liang, 1998; Sampson, 2000; Mathew & Mathew, 2001); POALES (Wodehouse, 1935; Erdtman, 1943, 1952; Skvarla & Larson, 1966; Christensen et al., 1972; Punt, 1975; Ladd, 1977; Nilsson et al., 1977; Köhler & Lange, 1979; Zavada, 1983; El-Ghazaly & Jensen, 1986, 1990; Chaturvedi et al., 1994, 1998; Chaturvedi & Datta, 2001; Dönmez & Işik, 2008; Mbagwu et al., 2008); PROTEALES (Wodehouse, 1935; Erdtman, 1943; Denk & Tekleva, 2006); RANUNCULALES (Erdtman, 1952; Nowicke & Skvarla, 1981; Clarke et al., 1991; Huysmans et al., 1998; Furness et al., 2007); ROSALES (Erdtman, 1952; Punt & Malotaux, 1984; Pehlivan, 1987); SAPINDALES (Erdtman, 1952; Grant et al., 2000); SOLANALES (Erdtman, 1952; Godwin et al., 1967; Godwin, 1968; Sengupta, 1972; Sharma, 1974; Punt & Monnabrands, 1977; Cresti et al., 1985; Trigo, 1992; Huysmans et al., 1998; Tellería & Daners, 2003; Al-Wadi & Lashin, 2007; Zhang et al., 2009); VITALES (Erdtman, 1952; Sharma, 1967; Zavada, 1983; Inceoglu et al., 2000; Punt et al., 2003a; Perveen & Qaiser, 2008b).

APPENDIX 2. Genera for which palynological data were not available for the specific taxa sampled by Jansen et al. (2007) in the molecular phylogeny, giving the closely related taxon used in their place, with angiosperms listed alphabetically by genus.

Gymnosperm outgroups: *Cycas revoluta* Thunb. data used for *C. micronesica* K. D. Hill; *Pinus scopulorum* (Engelm.) Lemmon and *P. sylvestris* L. for *P. thunbergii* Parl.

Angiosperm genera: *Acorus gramineus* Sol. ex Aiton data used for both *A. americanus* (Raf.) Raf. (Acorus.ame) and *A. calamus* L. (Acorus.cal); *Brassica campestris* L., *Brassica napus* L., and *Brassica tournefortii* Gouan for *Brassica rapa* L.; *Buxus sempervirens* L. for *Buxus microphylla* Siebold & Zucc.; *Campanula latifolia* L. and *Campanula strigosa* Vahl for *Trachelium caeruleum* L. (aff. *Campanula* L. sensu Roquet et al., 2008); *Chloranthus inconspicuus* Sw. and *Chloranthus japonicus* Siebold for *Chloranthus spicatus* (Thunb.) Makino; *Citrus limon* (L.) Osbeck and *Citrus medica* L. for *Citrus sinensis* (L.) Osbeck; *Coffea congensis* A. Froehner for *Coffea arabica* L.; *Cucumis melo* L. for *Cucumis sativus* L.; *Dioscorea polygonoides* Humb. & Bonpl. ex Willd. for *Dioscorea elephantipes* (L'Hér.) Engl.; *Drimys winteri* J. R. Forst. & G. Forst. for *Drimys granadensis* L. f.; *Elaeis guineensis* Jacq. for *E. oleifera* (Kunth) Cortés; *Illicium fargesii* Finet & Gagnep., *I. floridanum* J. Ellis, and *I. religiosum* Siebold & Zucc. for *I. oligandrum* Merr. & Chun; *Jasminum fruticans* L. and *J. humile* L. for *J. nudiflorum* Lindl.; *Medicago lupulina* L. for *Medicago truncatula* Gaertn.; *Morus alba* L. and *Morus nigra* L. for *Morus indica* L.; *Musa coccinea* Andrews for *Musa acuminata* Colla; *Nicotiana alata* Link & Otto, *Nicotiana glauca* Graham, and *Nicotiana tabacum* L. data used for all three species *Nicotiana sylvestris* Speg. & S. Comes (Nicotiana.syl), *Nicotiana tabacum* (Nicotiana.tab), and *Nicotiana tomentosa* (Nicotiana.tom), depending on the character (not all characters documented in the literature for all species); *Nymphaea colorata* Peter for *Nymphaea alba* L.; *Oenothera biennis* L. and *O. maysillesii* Munz for *O. elata* Kunth; *Passiflora caerulea* L. and *Passiflora suberosa* L. for *Passiflora biflora* Lam.; *Pelargonium grandiflorum* Willd. for *Pelargonium* × *hortorum* L. H. Bailey; *Phalaenopsis lueddemanniana* Rchb. f. (Orchidaceae) and data ascribed to other Vandeae for *Phalaenopsis aphrodite* Rchb. f.; *Piper majusculum* Blume (Piperaceae), *Piper nigrum* L., and *Piper polysyphonum* C. DC. for *Piper cenocladum* C. DC.; *Populus tremula* L. for *Populus trichocarpa* Torr. & A. Gray; *Ranunculus repens* L. for *R. macranthus* Scheele; *Scaevola taccada* (Gaertn.) Roxb. for *Scaevola aemula* R. Br.; *Solanum lycopersicum* L., *Solanum lysimachioides* Wall., and *Solanum macracanthum* A. Rich. data used for all three species *Solanum bulbocastaneum* Dunal (Solanum.bul), *Solanum lycopersicum* L. (Solanum.lyc), and *Solanum tuberosum* L. (Solanum.tub), depending on the character (not all characters documented in the literature for all species); *Vitis sylvestris* Blume for *V. vinifera* L.; *Yucca elephantipes* Hort. ex Regel for *Y. schidigera* Ortgies.

APPENDIX 3. List of pollen morphological characters and character states used in the four matrices studied. C, comprehensive method of coding; D, democratic method; E, exemplar method; E(NP), exemplar method with polymorphic data points treated as missing data.

1. Dispersal unit
   C: 0, monad; 1, dyad; 2, permanent tetrad; 3, polyad; 4, pollinia.
   D: 0, monad.
   E, E(NP): 0, monad; 1, permanent tetrad; 2, pollinia.
2. Polarity
   C, D, E, E(NP): 0, apolar; 1, heteropolar; 2, isopolar (including subisopolar).
3. Symmetry in polar view
   C, D, E, E(NP): : 0, bilateral; 1, radial.
4. Basic shape
   C, D, E, E(NP): 0, boat-shaped; 1, globose.
5. Shape class
   C: 0, peroblate; 1, oblate; 2, subspheroidal; 3, prolate; 4, perprolate.
   D: 0, oblate; 1, subspheroidal.
   E: 0, oblate; 1, subspheroidal; 2, prolate; 3, perprolate.
   E(NP): 0, oblate; 1, subspheroidal; 2, prolate.
6. Outline in polar view
   C: 0, elliptic; 1, circular; 2, polygonal; 3, lobate; 4, concave-polygonal.
   D, E, E(NP): 0, elliptic; 1, circular; 2, polygonal; 3, lobate.
7. Size
   C, E, E(NP): 0, very small; 1, small; 2, medium; 3, large; 4, very large.
   D: 0, small; 1, medium; 2, large.
8. Aperture number
   C, E, E(NP): 0, zero; 1, one; 2, two; 3, three; 4, four to six; 5, seven to 12; 6, many.
   D: 0, zero; 1, one; 2, three; 3, many.
9. Aperture position
   C, D, E, E(NP): 0, distal; 1, equatorial; 2, global.
10. Aperture structure
    C, D, E, E(NP): 0, simple; 1, compound.
11. Ectoaperture shape
    C: 0, porate; 1, colpate; 2, zonate; 3, spiral; 4, syncolpate.
    D: 0, porate; 1, colpate.
    E: 0, porate; 1, colpate; 2, spiral; 3, syncolpate.
    E(NP): 0, porate; 1, colpate; 2, syncolpate.
12. Aperture membrane ornamentation
    C, D, E, E(NP): 0, smooth; 1, granulate.
13. Operculum
    C, D, E, E(NP): 0, absent; 1, present.
14. Annulus
    C, E, E(NP): 0, absent; 1, costate; 2, aspidate; 3, vestibulate.
    D: 0, absent; 1, costate; 2, aspidate.
15. Exine differentiation
    C, D, E, E(NP): 0, absent; 1, present.
16. Supratectal elements
    C, D, E, E(NP): 0, absent; 1, present.
17. Supratectal element shape
    C, D: 0, verrucate; 1, echinate; 2, gemmate; 3, pilate.
    E, E(NP): 0, verrucate; 1, echinate; 2, gemmate.
18. Supratectal element size
    C, D, E, E(NP): 0, smaller than 1 μm; 1, greater than 1 μm.
19. Tectum sculpture
    C, E: 0, imperforate; 1, areolate; 2, *Croton*-pattern; 3, *Amborella*-type; 4, fossulate; 5, foveolate; 6, perforate; 7, reticulate; 8, rugulate; 9, striate; A, striato-reticulate.
    D: 0, imperforate; 1, *Croton*-patterned; 2, *Amborella*-type; 3, foveolate; 4, perforate; 5, reticulate; 6, striate.
    E(NP): 0, imperforate; 1, areolate; 2, *Croton*-patterned; 3, *Amborella*-type; 4, fossulate; 5, foveolate; 6, perforate; 7, reticulate; 8, rugulate; 9, striate.
20. Infratectum structure
    C, E, E(NP): 0, alveolate; 1, columellate; 2, granulate.
    D: 0, alveolate; 1, columellate.
21. Number of infratectum layers
    C, D, E, E(NP): 0, one; 1, two.
22. Foot layer
    C, D, E, E(NP): 0, absent; 1, present.
23. Endexine

C, D, E, E(NP): 0, absent; 1, present.
24. Endexine type
C, D, E, E(NP): 0, compact; 1, lamellar; 2, granular/spongy.
25. Endexine extent
C, D, E, E(NP): 0, continuous; 1, discontinuous; 2, apertural.
26. Cavea
C, D, E, E(NP): 0, absent; 1, present.
27. Internal foramina
C, D, E, E(NP): 0, absent; 1, present.
28. Microchannels
C, D, E, E(NP): 0, absent; 1, present.
29. Exine folding
C, D, E, E(NP): 0, non-lophate; 1, lophate.
30. Viscin threads
C, D, E, E(NP): 0, absent; 1, present.
31. Orbicules
C, D, E, E(NP): 0, absent; 1, present.

APPENDIX 4. Testing the impact of choice of priors for HB inference.

Multiple types of prior distribution were compared for two test characters, using the species exemplar matrix (polymorphic data points treated as missing data): Character 7 (pollen grain size), a multi-state character expected to display little structure with respect to the phylogenetic hypothesis; and Character 10 (aperture structure), a binary character expected to show clear structure with respect to the phylogenetic hypothesis. The priors tested were: uniform distribution; beta, exponential, and gamma distributions with parameters as indicated by the transition rates from the empirical Bayesian analysis; and a distribution generated by the reversible-jump hyperprior (rjhp) mechanism, as recommended in the BayesTraits manual (Pagel & Meade, 2006).

With both characters, the uniform prior distribution produced clearly unrealistic posterior distributions of inferred ancestral states, in which all states were almost equally probable at every internal node. This implies that a uniform prior distribution can dominate the results of small datasets, as suggested by Pagel and Meade (2006). Furthermore, likelihood values (-Lh) were much lower (e.g., for Character 7, −79.8 compared with −46 to −50) than with alternative prior distributions for the same data. Priors set using the rjhp, and with beta, exponential, and gamma priors using parameters derived from the EB analysis, all produced similar results in terms of inferred ancestral states.

For Character 7 (pollen grain size), the posterior distributions obtained using gamma- and exponentially distributed priors were most similar to one another; the posterior distributions obtained using the rjhp method were similar to those seen using EB; those obtained using beta-distributed priors were different from the other methods and tended to produce more ambiguous results, rather similar to using a uniform prior. The highest -Lh value was obtained using the rjhp approach. For Character 10 (aperture structure), the rjhp mechanism produced results most similar to those obtained with MP and ML, while the beta, exponential, and gamma prior distributions produced results very similar to one another, with slightly lower -Lh values than the rjhp mechanism.

Based on these results, in which beta-, exponential-, and gamma-distributed priors, as well as priors generated using the rjhp, produced broadly similar results, we chose to use the rjhp mechanism as recommended by Pagel and Meade (2006), specifying an exponential or gamma prior distribution where this suits the data for each particular character. The optimal parameters obtained for HB (continuous-time Markov model; aiming for an average acceptance rate of ca. 30% during rjhp optimization) ranged from gamma to exponential distributions, with seed means from 0 or 1 (Character 13) to 0 to 150 (Character 27), seed variances (for gamma distributions only) from 0 to 10 (Character 2) to 0 to 100 (Characters 19 and 26), and rate deviations from 5 (Character 15) to 180 (Character 17).

Appendix 5. Morphological matrices analyzed in this study (for simplicity, taxon names used are the same as those in Jansen et al. [2007] throughout).

Comprehensive method

Acorus.ame	0100-01100110010--{56}1011100??000
Acorus.cal	0100-01100110010--{56}1011100??000
Amborella	010{01}2{01}{12}{01}00{01}{01}0011103101{01}02000000
Anethum	02{01}1{234}{12}{123}{234}11110{012}10--{6789A}10110000000{01}
Arabidopsis	0{02}{01}1{123}{123}{1234}{0234}1{01}{14}{01}{01}{02}1{01}{01} 0{67A}101{01}10000000
Atropa	{02}{02}11{123}{0123}{1234}{3456}{12}{01}{01}{01}0{023}1{01}{0123}{01} {0567A}10111000000{01}
Brassica	0{02}{01}1{123}{123}{123}{0234}1{01}{14}{01}{01}{02} 1{01}{01}0{67A}101{01}10000000
Buxus	0{02}1121{12}{36}{12}{01}{01}0001{01}{01}0{27}101110000000
Calycanthus	{02}{02}{01}121{123}{02}1011001{01}{012}{01}{679}101111000000
Chloranthus	0{012}{01}{01}2{13}{12}{014}{01}011{01}010--7101110000000
Citrus	{02}{12}11{01234}{12}{123}{2345}1{01}{014}1{01} {012}1{01}{123}{01}{079A}1011?00??000
Coffea	{02}{12}{01}1{0123}{12}{1234}{012346}{12}{01}{0123}{01}0{02}1{01} {012}{01}{45679A}1011?00??000
Cucumis	{02}2{01}1{01234}{0123}{1234}{2346}1{01}{014} 10{023}1{01}11{5679}1011??0??000
Cycas	0100-{01}{123}100100010--{456}001110000000
Daucus	02{01}1{234}{12}{123}{234}11110{012}10--{6789A}10110000000{01}
Dioscorea	01{01}{01}-{01}{23}{1234}{01}0{01}1001{01}10{67}101120000000
Drimys	{02}1{01}11{12}2100{01}00{02}10--{67}101110000000
Elaeis	01001{02}{123}{12}{01}0{01}0{01}01{01}{01}{01}{67}101???0??000
Eucalyptus	{02}2{01}1{0123}{02}{01234}{234}1{01}{0134}{01}0{02}1{01}{02}0{01689}10 {01}1000000{01}0
Ginkgo	010{01}202100{01}0001{01}10{89}00111000?000
Glycine	{023}211{0123}{123}{1234}{23456}111{01}0{013}1 {01}10{0479A}10{01}1?0000000
Gossypium	{02}{02}{01}1{123}{12}{1234}{23456}{12}{01}{014}1{01}{012}1 {01}{01}{01}{267A}1011{02}{01}000000
Helianthus	{02}211{0123}{12}{1234}{2345}1{01}{014}{01}{01}{02}1{01}1{01} {045679}1{01}{01}1{01}0 {01}{01}0{01}00
Hordeum	{02}{012}{01}1{123}{12}{1234}{012346}00{0123}{01} {01}{12}1{01}{01}{01}{15789} 101{01}2000100{01}
Illicium	0{12}{01}1{12}{12}{123}{12345}{01}0{014}{01}001120{57}101100000000
Ipomoea	{02}{02}11{123}{0123}{1234}{3456}{12}{01}{01}{01}0{023}1{01}{0123}{01} {05679A}10111000000{01}
Jasminum	{02}{02}{01}1{01234}{123}{1234}{0123456}{12}{01}{01234} {01}{01}{02}1{01}{01}{01} {1679}10{01}1?00??000
Lactuca	{02}211{0123}{12}{1234}{2345}1{01}{014}{01}{01}{02}1{01}1{01}{045679}1 {01}{01}1{01}0{01}{01}0{01}00
Liriodendron	{02}{012}{01}{01}20{234}{0123}{01}0{12}1001{01}00{678}101110000000
Lotus	{023}211{0123}{123}{1234}{23456}111{01}0{013} 1{01}10{04789A}10{01}1?0000000
Manihot	{012}{02}11{01234}{0123}{1234}{03456}{12}{01}{0124}1 {01}{02}1{01}{012}{01}{02679}10111{01}000000
Medicago	{023}211{0123}{123}{1234}{23456}111{01}0{013} 1{01}10{04789A}10{01}1?0000000
Morus	{02}2{01}1{123}{01234}{123}{2345}1{01}{014} 1{01}{02}1{01}1{01}{1789}1011?0000000
Musa	0{01}1121{234}{0156}{02}0{01}?000{01}{12}{01}--------??000
Nandina	{02}{012}11{234}{123}{123}{013456}1{01}{013}{01} {01}01{01}{01}{01}{679A}10111{02}00000{01}

Appendix 5. Continued.

Nicotiana.syl	{02}{02}11{123}{0123}{1234}{3456}{12}{01}{01}{01}0{023}1{01}{0123}{01}{05679A}10111000000{01}
Nicotiana.tab	{02}{02}11{123}{0123}{1234}{3456}{12}{01}{01}{01}0{023}1{01}{0123}{01}{05679A}10111000000{01}
Nicotiana.tom	{02}{02}11{123}{0123}{1234}{3456}{12}{01}{01}{01}0{023}1{01}{0123}{01}{05679A}10111000000{01}
Nuphar	{02}10{01}2{012}{123}{01234}{01}0{12}1{01}01{01}{012}{01}{069}{12}01{01}1?000000
Nymphaea	{02}10{01}2{012}{123}{01234}{01}0{12}1{01}01{01}{012}{01}{069}{12}01{01}1?000000
Oenothera	{02}2{01}1{0123}{02}{01234}{234}1{01}{0134}{01}0{02}1{01}{02}0{01689}10{01}1000000{01}0
Oryza	{02}{012}{01}1{123}{12}{1234}{012346}00{0123}{01}{01}{12}1{01}{01}{01}{15789}101{01}2000100{01}
Panax	02{01}1{234}{12}{123}{234}11110{012}10--{6789A}10110000000{01}
Passiflora	{012}{02}11{01234}{0123}{1234}{03456}{12}{01}{0124}1{01}{02}1{01}{012}{01}{02679}10111{01}000000
Pelargonium	0{02}11{23}{123}{234}{036}{12}{01}{01}{01}001{01}{01}{01}{79A}101{01}?0001000
Phalaenopsis	{024}{012}{01}{01}2{01}{01234}{01236}00{013}1{01}01{01}{01}0{05679}1001?0000000
Pinus	0{01}{01}{01}2{01}{123}{01}00{01}{01}001{01}{12}0{1567}{02}01110000000
Piper	{02}1012{01}{0123}{0123}00{01}1001{01}{01}1{178}{02}011100??000
Platanus	02{01}1{012}{0124}{123}{234}1{01}{01}1001{01}{01}{01}{57}101100000000
Populus	{012}{02}11{01234}{0123}{1234}{03456}{12}{01}{0124}1{01}{02}1{01}{012}{01}{02679}10111{01}000000
Ranunculus	{02}{012}11{234}{123}{123}{013456}1{01}{013}{01}{01}01{01}{01}{01}{679A}10111{02}00000{01}
Saccharum	{02}{012}{01}1{123}{12}{1234}{012346}00{0123}{01}{01}{12}1{01}{01}{01}{15789}101{01}2000100{01}
Scaevola	{02}211{0123}{12}{1234}{2345}1{01}{014}{01}{01}{02}1{01}1{01}{045679}1{01}{01}1{01}0{01}{01}0{01}00
Solanum.bul	{02}{02}11{123}{0123}{1234}{3456}{12}{01}{01}{01}0{023}1{01}{0123}{01}{05679A}10111000000{01}
Solanum.lyc	{02}{02}11{123}{0123}{1234}{3456}{12}{01}{01}{01}0{023}1{01}{0123}{01}{05679A}10111000000{01}
Solanum.tub	{02}{02}11{123}{0123}{1234}{3456}{12}{01}{01}{01}0{023}1{01}{0123}{01}{05679A}10111000000{01}
Sorghum	{02}{012}{01}1{123}{12}{1234}{012346}00{0123}{01}{01}{12}1{01}{01}{01}{15789}101{01}2000100{01}
Spinacia	{02}{02}11{123}{12}{1234}{3456}{12}{01}{014}1{01}{02}1{01}{01}0{567}1011??0{01}?000
Trachelium	{02}211{0123}{12}{1234}{2345}1{01}{014}{01}{01}{02}1{01}1{01}{045679}1{01}{01}1{01}0{01}{01}0{01}00
Triticum	{02}{012}{01}1{123}{12}{1234}{012346}00{0123}{01}{01}{12}1{01}{01}{01}{15789}101{01}2000100{01}
Typha	{02}{012}{01}1{123}{12}{1234}{012346}00{0123}{01}{01}{12}1{01}{01}{01}{15789}101{01}2000100{01}
Vitis	0211{23}{123}{12}{34}11110110--{78}1011??000000
Yucca	{024}{012}{01}{01}2{01}{01234}{01236}00{013}1{01}01{01}{01}0{05679}1001?0000000
Zea	{02}{012}{01}1{123}{12}{1234}{012346}00{0123}{01}{01}{12}1{01}{01}{01}{15789}101{01}2000100{01}

Appendix 5. Continued.

Democratic method

Acorus.ame	0100-00100110010--?1010100??000
Acorus.cal	0100-00100110010--?1010100??000
Amborella	010?1??100??00111021O1?02000000
Anethum	021112121111O010--5101000000000
Arabidopsis	02111312101100l0?05101010000000
Atropa	021111121011001110410101O000000
Brassica	02111312101100l0?05101010000000
Buxus	00111113200000l0?05101010000000
Calycanthus	00111110101100111051010110000000
Chloranthus	02?11?11?0110010--5101010000000
Citrus	021111121111011030510l0?00??000
Coffea	02111112111l0?102051010?00??000
Cucumis	02111102111l0?1011510l0??0??000
Cycas	0100-01100100010--3001010000000
Daucus	021112121111O010--5101000000000
Dioscorea	0100-01100110010104101020000000
Drimys	0101011100100?10--?101010000000
Elaeis	0100001100100O1?105101???0??000
Eucalyptus	021112021111021000010100000000O
Ginkgo	010?101100?0001?10?00101000?000
Glycine	02111112111100101061000?0000000
Gossypium	00111123201l?2111111010?0000000
Helianthus	02111112111100111141?10?0000000
Hordeum	011111110000121100510102000l000
Illicium	0111011?10??00112051010000000O
Ipomoea	02111112101100111041010l0000000
Jasminum	02111112101100101051010?00??000
Lactuca	02111112111100111141?10?0000000
Liriodendron	0100102100110O1?00510101O0000000
Lotus	02111112111100101061000?0000000
Manihot	02111112111100101051010l?000000
Medicago	02111112111100101061000?0000000
Morus	02111112111102101061010?0000000
Musa	00111120?00?000110--------??000
Nandina	02111112101100111051010l0000000?
Nicotiana.syl	02111112101100111041010l0000000
Nicotiana.tab	02111112101100111041010l0000000
Nicotiana.tom	02111112101100111041010l0000000
Nuphar	010?10?10011?0111??10101?000000
Nymphaea	010?10?10011?0111??10101?000000
Oenothera	02111202111102100001010000000O
Oryza	011111110000121100510102000l000
Panax	021112121111O010--5101000000000
Passiflora	02111112111100101051010l?000000
Pelargonium	02111122101100111051010?0001000
Phalaenopsis	01011111001100111051000?0000000
Pinus	011111110000001020?001010000000
Piper	01011101001100l001?10101O0??000
Platanus	02110212100100101051010000000O
Populus	02111112111100101051010l?000000
Ranunculus	02111112101100111051010l0000000?
Saccharum	011111110000121100510102000l000
Scaevola	02111112111100111141?10?0000000
Solanum.bul	02111112101100111041010l0000000
Solanum.lyc	02111112101100111041010l0000000
Solanum.tub	02111112101100111041010l0000000
Sorghum	011111110000121100510102000l000
Spinacia	02111112101100111041010??00?000

Appendix 5. Continued.

Trachelium	021111121111001111141?10?0000000
Triticum	011111110000121100510102000100 0
Typha	011111110000121100510102000100 0
Vitis	021112121111011 0--51010??000000
Yucca	010111110011001110510 00?0000000
Zea	011111110000121100510102000100 0
Exemplar species method	
Acorus.ame	0100-01100110010--610111 00??000
Acorus.cal	0100-01100110010--510111 00??000
Amborella	010{01}1{01}{12}{01}00{01}{01}0011103101{01}02000000
Anethum	0211{23}223111?0110--910????0??000
Arabidopsis	021113{12}310100010--7101110000000
Atropa	02111123111103110 0{79}10????0??000
Brassica	0211111310110010--71010--000000
Buxus	00111126200000 1{01}007101110000000
Calycanthus	020111221011 0010--7101111000000
Chloranthus	02111{13}1410110010--7101110000000
Citrus	021111241111011 0--71011?00??000
Coffea	0{12}11{12}12{34}{12}1{12}1001120{6479}1?????0??000
Cucumis	02110{12}3310010310--710????0??000
Cycas	0100-02100100010--4001110000000
Daucus	021122231111?0110--{89}10110000000{01}
Dioscorea	0100-03210110010--6101120000000
Drimys	11110{12}2100000210--7101110000000
Elaeis	0100-02100100010--7101???0??000
Eucalyptus	021102{12}31130021{01}00{068}101100000000
Ginkgo	010{01}102100{01}0001{01}10{89}00111000?000
Glycine	021111131110 0010--71???????000
Gossypium	001111{34}62001{01}{12}1111{67}10110000?000
Helianthus	02111123111100111161111110110000
Hordeum	0111112100001211 10{18}1011??0??000
Illicium	0211{01}123103000112071011 00000000
Ipomoea	001111{34}6200100111 1010?10000000{01}
Jasminum	021111{23}311110010--710????0??000
Lactuca	02111223111000 1111{67}1111?0000100
Liriodendron	0100--031001?001100{68}101110000000
Lotus	021121131110 0010--81001?00??000
Manihot	001111{34}6200?00112021011 1000000
Medicago	0211232{34}111?0010--{48}10????0??000
Morus	0211111{234}100112111081011??000000
Musa	00111140------00--0------0-?000
Nandina	021123231011 0010--{67}101110000001
Nicotiana.syl	02111123111 0010--{7A}10????000000
Nicotiana.tab	0211112{34}11110010--{7A}10????000000
Nicotiana.tom	0211112{34}11110010--{7A}10????000000
Nuphar	010{01}10{23}1001110111{01}?20111?00?000
Nymphaea	010{01}102110111011{012}1?20111?00?000
Oenothera	0211{01}24310010311({02}00100100000010
Oryza	0111112100001211 1011010--0??00{01}
Panax	02111{12}23111?0{01}10--{6A}1011?000?000
Passiflora	021111125103?0010--7101110000000
Pelargonium	02112123101{01}0010--7101{01}--001000
Phalaenopsis	20?11100------10--01001?0000000
Pinus	0100-03100100010--7001110000000
Piper	0101111100110011{01}18{12}011100??000
Platanus	021111{12}310110010--7101100000000
Populus	00111120------1120710111 0000000
Ranunculus	0211112{345}10110011106101 1?20??000

Appendix 5. Continued.

Saccharum	011111121000?12110071010--0??000
Scaevola	021121231111021110{06}1111?00??000
Solanum.bul	021112231111031120010111?0??001
Solanum.lyc	021112231111031120610111?0??001
Solanum.tub	021112231111031120610111?0??001
Sorghum	011111121000?12110011010--001000
Spinacia	001111262001001110610???0??000
Trachelium	021111231100?10111041011110000000
Triticum	011111{23}1000012111011011--001000
Typha	111111{123}1000{01}0110--71010--00?000
Vitis	0211{12}21{34}11110110--{78}1011??000000
Yucca	0100-0{23}100110010--610???000000
Zea	011111{34}10000121110710112000100{01}
Exemplar species method, polymorphic data points treated as missing data	
Acorus.ame	0100-01100110010--610111100??000
Acorus.cal	0100-01100110010--510111100??000
Amborella	010?1???00??00111031012?02000000
Anethum	0211?223111?0110--910????0??000
Arabidopsis	021113?310100010--7101110000000
Atropa	0211112311110311002?10????0??000
Brassica	021111131011001O--71010--000000
Buxus	0011112620000012?007101110000000
Calycanthus	020111221011001O--7101111000000
Chloranthus	02111?1410110010--7101110000000
Citrus	021111241111011O--71011?00??000
Coffea	0?11?12??1?1001120?1?????0??000
Cucumis	02110?3310010310--710????0??000
Cycas	0100-02100100010--40011100000000
Daucus	021122231112?0110--?10110000000?
Dioscorea	0100-03210110010--61011200000000
Drimys	11110?2100000210--7101110000000
Elaeis	0100-02100100010--7101???0??000
Eucalyptus	021102?31120021?00?1011?0000000
Ginkgo	010?102100?0001?10?00111000?000
Glycine	021111131110001O--71????????000
Gossypium	001111?62001??1111?10110000?000
Helianthus	021111231111001111161111110110000
Hordeum	011111210000121110?1011??0??000
Illicium	0211?123102000112071011000000000
Ipomoea	001111?62001001111010?100000000?
Jasminum	021111?311110010--710????0??000
Lactuca	021112231110001111?1111?0000100
Liriodendron	0100--031001?001100?101110000000
Lotus	02112113111000l0--81001?00??000
Manihot	001111?6200?001120210111l000000
Medicago	0211232?111?0010--?10????0??000
Morus	0211111?100112111081011??000000
Musa	00111140------00--0------0-?000
Nandina	021123231011001O--?101110000001
Nicotiana.syl	02111123111100l0--?10????000000
Nicotiana.tab	0211112?11110010--?10????000000
Nicotiana.tom	0211112?11110010--?10????000000
Nuphar	010?10?100111011l??20111?00?000
Nymphaea	010?102110111011?1?20111?00?000
Oenothera	0211?2431001031l?001001000000l0
Oryza	011111210000121110l1010--0??00?
Panax	02111?23111?0?10--?1011?000?000
Passiflora	021111121111001010510101?000000
Pelargonium	02112123101?0010--7101?--001000

Appendix 5. Continued.

Phalaenopsis	20?11100------10--01001?00000000
Pinus	0100-03100100010--70011100000000
Piper	0101111100110011?18?011100??000
Platanus	021111?310110010--71011000000000
Populus	00111120------1120710111000000
Ranunculus	0211112?10110011106101l?20??000
Saccharum	011111210000?12110071010--0??000
Scaevola	021121231111021110?1111?00??000
Solanum.bul	0211122311110311200l0111?0??001
Solanum.lyc	021112231111031120610111?0??001
Solanum.tub	021112231111031120610111?0??001
Sorghum	011111210000?12110011010--001000
Spinacia	00111126200100111061O????0??000
Trachelium	021111231OO?10111041011100000000
Triticum	011111?1000012111011011--001000
Typha	111111?1000?0110--71010--00?000
Vitis	0211?21?11110110--?1011??000000
Yucca	0100-0?100110010--610????000000
Zea	011111?10000121110710112000100?

Appendix 6. Voucher details for pollen grains shown in Figure 3.

Taxon	Collection	Locality	Herbarium	Figure
Acacia nilotica (L.) Willd. ex Delile (Fabaceae)	*J. B. Allen 229*	Afgoi, Somalia	E	4C
Acorus gramineus Sol. ex Aiton (Acoraceae)	*K. L. Chu 2272*	Tien Chuan Hsien, China	E	4E
Adenocaulon chilense Less. (Asteraceae)	*T. M. Pedersen 1494*	Los Lagos, Argentina	US	4R
Aesculus hippocastanum L. (Sapindaceae)	—	cult. nr. Haywards Heath, United Kingdom	—	4P
Alfredia cernua (L.) Cass. (Asteraceae)	*Roginskaya 74*	Russia	K	4X
Cabobanthus bullulatus (S. Moore) H. Rob. (Asteraceae)	*Lisowski 357*	Haut-Katanga, Democratic Republic of Congo	K	4I
Centrapalus pauciflorus (Willd.) H. Rob. (Asteraceae)	*N. Hepper 3760*	Kati, Mali	K	4G
Crepis napifera (Franch.) Babc. (Asteraceae)	*G. Forrest 2987*	Yunnan, China	BM	4D
Croton argyratus Blume (Euphorbiaceae)	*B. S. Expedition 201*	Thailand	E	4U, V
Cullumia rigida DC. (Asteraceae)	*T. Trinder-Smith 182*	South Africa	US	4T
Dampiera stricta (Sm.) R. Br. (Goodeniaceae)	*W. E. Fisher 99*	New South Wales, Australia	E	4W
Dioscorea nipponica Makino (Dioscoreaceae)	*O. Zhurban 6856*	Primorje, Russia	E	4F
Dioscorea pyrenaica Bubani & Bordere ex Gren. (Dioscoreaceae)	*C. Packe s.n.*	Spanish Pyrenees, Spain	E	4Q
Illicium floridanum J. Ellis (Illiciaceae)	*F. H. Utech 83-039*	Louisiana, U.S.A.	KUN	4O
Lathyrus pratensis L. (Fabaceae)	*Th. Constantinidis 8046*	Mount Kratsovon, Thessalia, Greece	E	4S
Nouelia insignis Franch. (Asteraceae)	*J. Rock 11714*	Yunnan, China	US	4J
Oldenburgia paradoxa Less. (Asteraceae)	*Owan 374*	South Africa	US	4L
Plantago psyllium L. (Plantaginaceae)	*Davis 52297*	Algeria	E	4A
Poa bulbosa L. (Poaceae)	*D. Davis 63473*	Agrigento, Sicily, Italy	E	4M
Rhododendron wallichii Hook. f. (Ericaceae)	*R. C. Poudel 10*	Royal Botanic Garden Edinburgh	E	4B
Scorzonera hispanica L. (Asteraceae)	—	cult. Chelsea Physic Garden, United Kingdom	—	4N
Tragopogon longifolius Heldr. & Sartori (Asteraceae)	*F. Guiol 1007*	Attica, Greece	BM	4K
Ulmus glabra Huds. (Ulmaceae)	*Culte 15278*	Royal Botanic Garden Edinburgh	E	4H